

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2002, 08:01:57 ; Search time 70.27 Seconds
(without alignments)
1824.686 Million cell updates/sec

Title: US-09-680-514-4_COPY_526_1047

Perfect score: 522
Sequence: 1 GCGCCACATATGCGCCTC.....TAGGCCACCTTGCCAGCCC 522

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*

2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*

3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	486.8	93.3	525	1	US-08-434-411-1
2	486.8	93.3	525	1	US-08-434-402-1
3	486.8	93.3	525	1	US-08-783-288-1
4	486.8	93.3	525	2	US-08-890-640-1
5	486.8	93.3	525	6	5194592-25
6	474.6	90.9	546	3	US-08-469-318-177
7	474.6	90.9	546	3	US-08-469-318-178
8	474.6	90.9	546	3	US-08-468-609A-177
9	474.6	90.9	546	3	US-08-468-609A-178
10	474.6	90.9	546	5	PCT-US95-01185-177
11	474.6	90.9	546	5	PCT-US95-01185-178
12	474.6	90.9	921	3	US-08-469-318-72
13	474.6	90.9	921	3	US-08-469-318-75
14	474.6	90.9	921	3	US-08-469-318-78
15	474.6	90.9	921	3	US-08-469-318-84
16	474.6	90.9	921	3	US-08-468-609A-72
17	474.6	90.9	921	3	US-08-468-609A-75
18	474.6	90.9	921	3	US-08-468-609A-78
19	474.6	90.9	921	3	US-08-468-609A-84
20	474.6	90.9	921	5	PCT-US95-01185-72
21	474.6	90.9	921	5	PCT-US95-01185-75
22	474.6	90.9	921	5	PCT-US95-01185-78
23	474.6	90.9	921	5	PCT-US95-01185-84
24	474.6	90.9	966	3	US-08-469-318-68
25	474.6	90.9	966	3	US-08-469-318-70
26	474.6	90.9	966	3	US-08-469-318-71
27	474.6	90.9	966	3	US-08-469-318-73

28	474.6	90.9	966	3	US-08-469-318-77
29	474.6	90.9	966	3	US-08-469-318-79
30	474.6	90.9	966	3	US-08-468-609A-68
31	474.6	90.9	966	3	US-08-468-609A-70
32	474.6	90.9	966	3	US-08-468-609A-71
33	474.6	90.9	966	3	US-08-468-609A-73
34	474.6	90.9	966	3	US-08-468-609A-77
35	474.6	90.9	966	3	US-08-468-609A-79
36	474.6	90.9	966	5	PCT-US95-01185-68
37	474.6	90.9	966	5	PCT-US95-01185-70
38	474.6	90.9	966	5	PCT-US95-01185-71
39	474.6	90.9	966	5	PCT-US95-01185-73
40	474.6	90.9	966	5	PCT-US95-01185-77
41	474.6	90.9	966	5	PCT-US95-01185-79
42	474.6	90.9	1017	3	US-08-469-318-67
43	474.6	90.9	1017	3	US-08-468-609A-67
44	474.6	90.9	1017	5	PCT-US95-01185-67
45	474.6	90.9	1047	3	US-08-469-318-74

ALIGNMENTS

RESULT 1
US-08-434-411-1
; Sequence 1, Application US/08434411
; Patent No. 5681720
; GENERAL INFORMATION:
; APPLICANT: KUGA, TETSURO
; APPLICANT: MIYAJI, HIROMASA
; APPLICANT: SATO, MORIYUKI
; APPLICANT: OKABE, MASAMI
; APPLICANT: MORIMOTO, MAKOTO
; APPLICANT: ITOH, SEIGA
; APPLICANT: YAMASAKI, MOTOO
; APPLICANT: YOKOO, YOSHIHARU
; APPLICANT: YAMAGUCHI, KAZUO
; APPLICANT: YOSHIDA, HAJIME
; APPLICANT: YOSHINORI, KOMATSU
; TITLE OF INVENTION: NOVEL POLYPEPTIDES
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,411
; FILING DATE: 03-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 306799/86
; FILING DATE: 23-DEC-1986
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 51357/88
; FILING DATE: 04-MAR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 80088/88
; FILING DATE: 31-MAR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAWFORD, ARTHUR
; REGISTRATION NUMBER: 25327
; REFERENCE/DOCKET NUMBER: 249-73
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4000
; TELEFAX: (703)816-4100

```
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 525 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..522
; OTHER INFORMATION: /product= "HUMAN GRANULOCYTE COLONY
; OTHER INFORMATION: STIMULATING FACTOR"
US-08-434-411-1

Query Match 93.3%; Score 486.8; DB 1; Length 525;
Best Local Similarity 98.6%; Pred. No. 2.3e-109;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 25 CTACCACAGAGCTTCCTTTTAAAGCTTAGAGCAAGTGAAGAGATCCAGGGCGATGGC 84
DB 25 CTGCCCCAGAGCTTCCTGCTCAAGTGTAGAGCAAGTGAAGAGATCCAGGGCGATGGC 84
QY 85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAAGCTGTGCCACCCCGAGGAGCTGGT 144
DB 85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAAGCTGTGCCACCCCGAGGAGCTGGT 144
QY 145 CTGCTCGAGACTCTCTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCCCAGCCAGGCC 204
DB 145 CTGCTCGAGACTCTCTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCCCAGCCAGGCC 204
QY 205 CTGACGTGGCAGGCTGCTTAAAGCTTAGAGCAAGTGAAGAGATCCAGGGCGATGGC 264
DB 205 CTGACGTGGCAGGCTGCTTAAAGCTTAGAGCAAGTGAAGAGATCCAGGGCGATGGC 264
QY 265 CTGACGGCCCTGGAAGGATCTCCCCCGAGTTGGTGGGCTTGGACACACTGCAGCTG 324
DB 265 CTGACGGCCCTGGAAGGATCTCCCCCGAGTTGGTGGGCTTGGACACACTGCAGCTG 324

Query Match 93.3%; Score 486.8; DB 1; Length 525;
Best Local Similarity 98.6%; Pred. No. 2.3e-109;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 25 CTACCACAGAGCTTCCTTTTAAAGCTTAGAGCAAGTGAAGAGATCCAGGGCGATGGC 84
DB 25 CTGCCCCAGAGCTTCCTGCTCAAGTGTAGAGCAAGTGAAGAGATCCAGGGCGATGGC 84
QY 85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAAGCTGTGCCACCCCGAGGAGCTGGT 144
DB 85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAAGCTGTGCCACCCCGAGGAGCTGGT 144
QY 145 CTGCTCGAGACTCTCTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCCCAGCCAGGCC 204
DB 145 CTGCTCGAGACTCTCTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCCCAGCCAGGCC 204
QY 205 CTGACGTGGCAGGCTGCTTAAAGCTTAGAGCAAGTGAAGAGATCCAGGGCGATGGC 264
DB 205 CTGACGTGGCAGGCTGCTTAAAGCTTAGAGCAAGTGAAGAGATCCAGGGCGATGGC 264
QY 265 CTGACGGCCCTGGAAGGATCTCCCCCGAGTTGGTGGGCTTGGACACACTGCAGCTG 324
DB 265 CTGACGGCCCTGGAAGGATCTCCCCCGAGTTGGTGGGCTTGGACACACTGCAGCTG 324
```

RESULT 2

US-08-434-402-1

; Sequence 1, Application US/08434402

; Patent No. 5714581

; GENERAL INFORMATION:

; APPLICANT: KUGA, TETSURO

; APPLICANT: MIYAJI, HIROMASA

; APPLICANT: SATO, MORIYUKI

; APPLICANT: OKABE, MASAMI

; APPLICANT: MORIMOTO, MAKOTO

; APPLICANT: ITOH, SEIGA

; APPLICANT: YAMASAKI, MOTOO

; APPLICANT: YOKOO, YOSHIHARU

; APPLICANT: YAMAGUCHI, KAZUO

; APPLICANT: YOSHIDA, HAJIME

; APPLICANT: YOSHINORI, KOMATSU

```
; TITLE OF INVENTION: NOVEL POLYPEPTIDES
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,402
; FILING DATE: 03-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 306799/86
; FILING DATE: 23-DEC-1986
; APPLICATION NUMBER: JP 51357/88
; FILING DATE: 04-MAR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 80088/88
; FILING DATE: 31-MAR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAWFORD, ARTHUR
; REGISTRATION NUMBER: 25327
; REFERENCE/DOCKET NUMBER: 249-72
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4000
; TELEFAX: (703)816-4100
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 525 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..522
; OTHER INFORMATION: /product= "HUMAN GRANULOCYTE COLONY
; OTHER INFORMATION: STIMULATING FACTOR"
US-08-434-402-1
```

Query Match 93.3%; Score 486.8; DB 1; Length 525;
Best Local Similarity 98.6%; Pred. No. 2.3e-109;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
QY 25 CTACCACAGAGCTTCCTTTTAAAGCTTAGAGCAAGTGAAGAGATCCAGGGCGATGGC 84
DB 25 CTGCCCCAGAGCTTCCTGCTCAAGTGTAGAGCAAGTGAAGAGATCCAGGGCGATGGC 84
QY 85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAAGCTGTGCCACCCCGAGGAGCTGGT 144
DB 85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAAGCTGTGCCACCCCGAGGAGCTGGT 144
QY 145 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCCCAGCCAGGCC 204
DB 145 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCCCAGCCAGGCC 204
QY 205 CTGACGTGGCAGGCTGCTTAAAGCTTAGAGCAAGTGAAGAGATCCAGGGCGATGGC 264
DB 205 CTGACGTGGCAGGCTGCTTAAAGCTTAGAGCAAGTGAAGAGATCCAGGGCGATGGC 264
QY 265 CTGACGGCCCTGGAAGGATCTCCCCCGAGTTGGTGGGCTTGGACACACTGCAGCTG 324
DB 265 CTGACGGCCCTGGAAGGATCTCCCCCGAGTTGGTGGGCTTGGACACACTGCAGCTG 324
```


Db 505 cgccacccctggccagccc 522
|||||

RESULT 6
US-08-469-318-177
; Sequence 177, Application US/08469318
; Patent No. 6022535
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 196
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,318
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/446,872
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 177:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 546 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-469-318-177

Query Match 90.9%; Score 474.6; DB 3; Length 546;
Best Local Similarity 94.4%; Pred. No. 2.le-106;
Matches 492; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY 2 CGCCCAACATATCGCGCTCGAGCTTACCACAGAGCTTCTTTTAAAGAGCTTAGAGCAAG 61
DB 8 CACCATTAGGCTGCCAGCTCCCTGCCAGAGCTTCTGCTCAAGTCTTAGAGCAAG 67
QY 62 TGAGGAAGATCCAGGGCGATGCGCAGCGCTCCAGGAGAAGCTGTGTGCCACTTACAAGC 121
DB 68 TGAGGAAGATCCAGGGCGATGCGCAGCGCTCCAGGAGAAGCTGTGTGCCACTTACAAGC 127
QY 122 TGTGCCACCCCGAGAGCTGGTGTCTCGGACACTCTCTGGGCATCCCTGGGCTCCCC 181
DB 128 TGTGCCACCCCGAGAGCTGGTGTCTCGGACACTCTCTGGGCATCCCTGGGCTCCCC 187
QY 182 TGAGCAGCTGCCCGCAGCGCCCTGCAGCTGGCAGGCTGTGTAGCCCAACTCCATAGCG 241
DB 188 TGAGCTCTGCCCGCAGCGCCCTGCAGCTGGCAGGCTGTGTAGCCCAACTCCATAGCG 247
QY 242 GCCTTTTCTCTACAGGGGCTCTCTGAGGGCCCTTGAAGGGATCTCCCGGAGTTGGGTC 301
DB 248 GCCTTTTCTCTACAGGGGCTCTCTGAGGGCCCTTGAAGGGATATCCCGGAGTTGGGTC 307
QY 302 CCACCTTGGACACACTGCAGCTGGAGCTGCGGACTTTGCCACCACTCTCTGGCAGCAGA 361
DB 308 CCACCTTGGACACACTGCAGCTGGAGCTGCGGACTTTGCCACCACTCTCTGGCAGCAGA 367
QY 362 TGAAGAAGCTGGGAATGCCCTGCCCTGCAGCCACCCAGGGTGCCTATGCCGGCTTCG 421
DB 368 TGAAGAAGCTGGGAATGCCCTGCCCTGCAGCCACCCAGGGTGCCTATGCCGGCTTCG 427
QY 422 CCTCTGCTTTCAGCGCGGCGAGGGGTCCTAGTTGCCCTCCCATCTGCAGAGCTTCC 481
DB 428 CCTCTGCTTTCAGCGCGGCGAGGGGTCCTAGTTGCCCTCCCATCTGCAGAGCTTCC 487
QY 482 TGGAGGTGCTGACCGGCTTCTACGCCACCTTGGCCAGGCC 522
|||||

Db 488 TGGAGGTGCTGACCGGCTTCTACGCCACCTTGGCCAGGCC 528

RESULT 7
US-08-469-318-178
; Sequence 178, Application US/08469318
; Patent No. 6022535
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 196
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,318
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/446,872
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 178:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 546 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-469-318-178

Query Match 90.9%; Score 474.6; DB 3; Length 546;
Best Local Similarity 94.4%; Pred. No. 2.le-106;
Matches 492; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY 2 CGCCCAACATATCGCGCTCGAGCTTACCACAGAGCTTCTTTTAAAGAGCTTAGAGCAAG 61
DB 8 CACCATTAGGCTGCCAGCTCCCTGCCAGAGCTTCTGCTCAAGTCTTAGAGCAAG 67
QY 62 TGAGGAAGATCCAGGGCGATGCGCAGCGCTCCAGGAGAAGCTGTGTGCCACTTACAAGC 121
DB 68 TGAGGAAGATCCAGGGCGATGCGCAGCGCTCCAGGAGAAGCTGTGTGCCACTTACAAGC 127
QY 122 TGTGCCACCCCGAGAGCTGGTGTCTCGGACACTCTCTGGGCATCCCTGGGCTCCCC 181
DB 128 TGTGCCACCCCGAGAGCTGGTGTCTCGGACACTCTCTGGGCATCCCTGGGCTCCCC 187
QY 182 TGAGCAGCTGCCCGCAGCGCCCTGCAGCTGGCAGGCTGTGTAGCCCAACTCCATAGCG 241
DB 188 TGAGCTCTGCCCGCAGCGCCCTGCAGCTGGCAGGCTGTGTAGCCCAACTCCATAGCG 247
QY 242 GCCTTTTCTCTACAGGGGCTCTCTGAGGGCCCTTGAAGGGATCTCCCGGAGTTGGGTC 301
DB 248 GCCTTTTCTCTACAGGGGCTCTCTGAGGGCCCTTGAAGGGATATCCCGGAGTTGGGTC 307
QY 302 CCACCTTGGACACACTGCAGCTGGAGCTGCGGACTTTGCCACCACTCTCTGGCAGCAGA 361
DB 308 CCACCTTGGACACACTGCAGCTGGAGCTGCGGACTTTGCCACCACTCTCTGGCAGCAGA 367
QY 362 TGAAGAAGCTGGGAATGCCCTGCCCTGCAGCCACCCAGGGTGCCTATGCCGGCTTCG 421
DB 368 TGAAGAAGCTGGGAATGCCCTGCCCTGCAGCCACCCAGGGTGCCTATGCCGGCTTCG 427
QY 422 CCTCTGCTTTCAGCGCGGCGAGGGGTCCTAGTTGCCCTCCCATCTGCAGAGCTTCC 481
DB 428 CCTCTGCTTTCAGCGCGGCGAGGGGTCCTAGTTGCCCTCCCATCTGCAGAGCTTCC 487
QY 482 TGGAGGTGCTGACCGGCTTCTACGCCACCTTGGCCAGGCC 522
|||||

TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-468-609A-178

Query Match 90.9%; Score 474.6; DB 3; Length 546;
Best Local Similarity 94.4%; Pred. No. 2.1e-106;
Matches 492; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 2 CGCCAAACATATCGCGCCTCGAGTCTACACAGAGCTTCTTTTAAAGCTTAGAGCAAG 61
Db 8 CACCATTAGGACCTGCCAGTCCCTGCCCCAGAGCTTCTGCTCAAGTGTCTTAGAGCAAG 67
Qy 62 TGAGGAAGATCCAGGGCGATGGCGAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAGC 121
Db 68 TGAGGAAGATCCAGGGCGATGGCGAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAGC 127
Qy 122 TGTGCCACCCCGAGAGCTGTGTGCTCGGACACTCTCTGGGCAATCCCTGGGCTCCCC 181
Db 128 TGTGCCACCCCGAGAGCTGTGTGCTCGGACACTCTCTGGGCAATCCCTGGGCTCCCC 187
Qy 182 TGAGCAGCTGCCCGCAGCGCCCTGACAGCTGGCAGGCTGTGAGCAACTTCCATAGCG 241
Db 188 TGAGCTCTGCCCCAGCGCCCTGACAGCTGGCAGGCTGTGAGCAACTTCCATAGCG 247
Qy 242 GCCTTTTCTCTACAGGGGCTCTGACAGGCTTGGAAAGGATCTCCCGGAGTTGGGTC 301
Db 248 GCCTTTTCTCTACAGGGGCTCTGACAGGCTTGGAAAGGATATCCCGCGAGTTGGGTC 307
Qy 302 CCACCTTGGACACACTGACAGCTGACGCTGCGGAGCTTTGGCCACCACCATCTGGCAGCAGA 361
Db 308 CCACCTTGGACACACTGACAGCTGACGCTGCGGAGCTTTGGCCACCACCATCTGGCAGCAGA 367
Qy 362 TGGAAAGACTGGGAATGCCCTGCGCTGACGCCACCCAGGCTGCCATGCCGCTTCG 421
Db 368 TGGAAAGACTGGGAATGCCCTGCGCTGACGCCACCCAGGCTGCCATGCCGCTTCG 427
Qy 422 CCTCTGCTTTCCAGCGCGGCGAGAGGGGTCCTAGTTGGCTTCCCATCTGCAGAGCTTCC 481
Db 428 CCTCTGCTTTCCAGCGCGGCGAGAGGGGTCCTAGTTGGCTTCCCATCTGCAGAGCTTCC 487
Qy 482 TGGAGGTGCTGACGGCTTCTAGCCACCTTGGCCAGCCC 522
Db 488 TGGAGGTGCTGACGGCTTCTAGCCACCTTGGCCAGCCC 528

RESULT 10

PCT-US95-01185-177

; Sequence 177, Application PC/TUS9501185

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion

; TITLE OF INVENTION: Protein

; NUMBER OF SEQUENCES: 196

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/01185

; FILING DATE: 02-FEB-1995

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/192325

; FILING DATE: 14-FEB-1994

; INFORMATION FOR SEQ ID NO: 177:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 546 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

PCT-US95-01185-177

Query Match 90.9%; Score 474.6; DB 5; Length 546;
Best Local Similarity 94.4%; Pred. No. 2.1e-106;
Matches 492; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 2 CGCCAAACATATCGCGCCTCGAGTCTACACAGAGCTTCTTTTAAAGCTTAGAGCAAG 61
Db 8 CACCATTAGGACCTGCCAGTCCCTGCCCCAGAGCTTCTGCTCAAGTGTCTTAGAGCAAG 67
Qy 62 TGAGGAAGATCCAGGGCGATGGCGAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAGC 121
Db 68 TGAGGAAGATCCAGGGCGATGGCGAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAGC 127
Qy 122 TGTGCCACCCCGAGAGCTGTGTGCTCGGACACTCTCTGGGCAATCCCTGGGCTCCCC 181
Db 128 TGTGCCACCCCGAGAGCTGTGTGCTCGGACACTCTCTGGGCAATCCCTGGGCTCCCC 187
Qy 182 TGAGCAGCTGCCCGCAGCGCCCTGACAGCTGGCAGGCTGTGAGCAACTTCCATAGCG 241
Db 188 TGAGCTCTGCCCCAGCGCCCTGACAGCTGGCAGGCTGTGAGCAACTTCCATAGCG 247
Qy 242 GCCTTTTCTCTACAGGGGCTCTGACAGGCTTGGAAAGGATCTCCCGGAGTTGGGTC 301
Db 248 GCCTTTTCTCTACAGGGGCTCTGACAGGCTTGGAAAGGATATCCCGCGAGTTGGGTC 307
Qy 302 CCACCTTGGACACACTGACAGCTGACGCTGCGGAGCTTTGGCCACCACCATCTGGCAGCAGA 361
Db 308 CCACCTTGGACACACTGACAGCTGACGCTGCGGAGCTTTGGCCACCACCATCTGGCAGCAGA 367
Qy 362 TGGAAAGACTGGGAATGCCCTGCGCTGACGCCACCCAGGCTGCCATGCCGCTTCG 421
Db 368 TGGAAAGACTGGGAATGCCCTGCGCTGACGCCACCCAGGCTGCCATGCCGCTTCG 427
Qy 422 CCTCTGCTTTCCAGCGCGGCGAGAGGGGTCCTAGTTGGCTTCCCATCTGCAGAGCTTCC 481
Db 428 CCTCTGCTTTCCAGCGCGGCGAGAGGGGTCCTAGTTGGCTTCCCATCTGCAGAGCTTCC 487
Qy 482 TGGAGGTGCTGACGGCTTCTAGCCACCTTGGCCAGCCC 522
Db 488 TGGAGGTGCTGACGGCTTCTAGCCACCTTGGCCAGCCC 528

RESULT 11

PCT-US95-01185-178

; Sequence 178, Application PC/TUS9501185

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion

; TITLE OF INVENTION: Protein

; NUMBER OF SEQUENCES: 196

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/01185

; FILING DATE: 02-FEB-1995

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/192325

; FILING DATE: 14-FEB-1994

; INFORMATION FOR SEQ ID NO: 178:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 546 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

PCT-US95-01185-178

Query Match	90.94%;	Score	474.6;	DB	5;	Length	546;
Best Local Similarity	94.4%;	Pred.	No. 2.1e-106;				
Matches	492;	Conservative	0;	Mismatches	29;	Indels	0;
Qy	2	CGCCACATATCGCGCTCGAGTCTACCCACAGAGCTTCCTTTTAAAGCTTAGAGCAAG	61				
Db	8	CACCATTAGAGCTTCCAGACTCCCTGCCCCAGAGCTTCCTGCTCAAGTGTCTTAGAGCAAG	67				
Qy	62	TGAGGAAGATCCAGGCGCATGGCGACGGCTCCAGGAGAGCTGTGTCACCTCAACAGC	121				
Db	68	TGAGGAAGATCCAGGCGCATGGCGACGGCTCCAGGAGAGCTGTGTGCCACCTACAAGC	127				
Qy	122	TGTGCCACCCCGAGGAGCTGGTGTGCTGGACACTCTCTGGGCATCCCTGGGGTCCCC	181				
Db	128	TGTGCCACCCCGAGGAGCTGGTGTGCTCGGACACTCTCTGGGCATCCCTGGGGTCCCC	187				
Qy	182	TGAGCAGCTGCCCCAGCGCCCTGCAGCTGGCAGGCTGCTTTGAGCCAACCTCATAGCG	241				
Db	188	TGAGCTCTCTGCCCCAGCGCCCTGCAGCTGGCAGGCTGCTTTGAGCCAACCTCATAGCG	247				
Qy	242	GCTTTTCTCTACAGGGGCTCTCTGAGGCGCTTGAAGGATCTCCCCGAGTTGGGTC	301				
Db	248	GCTTTTCTCTACAGGGGCTCTCTGAGGCGCTTGAAGGATATCCCCGAGTTGGGTC	307				
Qy	302	CCACCTTGGACACCTGCAGCTGCACCTGCCGACTTTGCCACACCACTTGGCAGCAGA	361				
Db	308	CCACCTTGGACACCTGCAGCTGCACCTGCCGACTTTGCCACACCACTTGGCAGCAGA	367				
Qy	362	TGGAAGACTGGGAATGCCCTGCCCTGCAGCCACCCAGGTTGCCATGCCGGCTTCG	421				
Db	368	TGGAAGACTGGGAATGCCCTGCCCTGCAGCCACCCAGGTTGCCATGCCGGCTTCG	427				
Qy	422	CCTCTGCTTTCAGCGCCGGCAGAGGGGTTCCTAGTTGCTTCCCTCATCTGCAGAGCTTCC	481				
Db	428	CCTCTGCTTTCAGCGCCGGCAGAGGGGTTCCTAGTTGCTTCCCTCATCTGCAGAGCTTCC	487				
Qy	482	TGAGGTGTCTACCGCGTTTCTAGCCACCTTGGCCAGGCC	522				
Db	488	TGAGGTGTCTACCGCGTTTCTAGCCACCTTGGCCAGGCC	528				

```

RESULT 12
US-08-469-318-72
; Sequence 72, Application US/08469318
; Patent No. 6022535
; GENERAL INFORMATION:
;
; APPLICANT: Multivariant IL-3 Hematopoiesis Fusion
;
; TITLE OF INVENTION: Protein
;
; NUMBER OF SEQUENCES: 196
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/469,318
;
; FILING DATE:
;
; CLASSIFICATION:
;
; PRIOR APPLICATION DATA:
;
; APPLICATION NUMBER: 08/446,872
;
; FILING DATE:
;
; INFORMATION FOR SEQ ID NO: 72:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 921 base pairs
;
; TYPE: nucleic acid
;
; STRANDEDNESS: double
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: DNA (genomic)
;
US-08-469-318-72

```

Query Match	90.9%;	Score 474.6;	DB 3;	Length 921;	
Best Local Similarity	94.4%;	Pred. No. 2.3e-106;			
Matches 492;	Conservative	0;	Mismatches 29;	Indels 0;	Gaps
Qy	2	CGCCAAACATATCGCGCTCGAGTCTACACAGAGCTTCCTTTTAAAGAGCTTAGACCAAG	61		
Db	8	CACCATTTAGGCCCTGCCAGGTCCTCGCCAGAGCTTCTTGCTCAAGTCTTAGACCAAG	67		
Qy	62	TGAGGAAGATCCAGGGCGATGGCGAGCGCTCCAGGAGAGCTGTGTGCCACCTACCAAGC	121		
Db	68	TGAGGAAGATCCAGGGCGATGGCGAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAGC	127		
Qy	122	TGTGCCACCCCGAGGAGCTGGTGTCTCGGACACTCTCTGGGCATCCCTCGGGCTCCCC	181		
Db	128	TGTGCCACCCCGAGGAGCTGGTGTCTCGGACACTCTCTGGGCATCCCTCGGGCTCCCC	187		
Qy	182	TGAGCAGCTGCCCCAGCCAGGCGCTCGAGCTGCGAGGCTGCTTGAGCCAACTCCATAGCG	241		
Db	188	TGAGCTCCTGCCCCAGCCAGGCGCTCGAGCTGCGAGGCTGCTTGAGCCAACTCCATAGCG	247		
Qy	242	GCCTTTTCTCTACCAAGGGCTCCTCGAGGCGCTTGAAGGGATCTCCCGCAGTTGGGTC	301		
Db	248	GCCTTTTCTCTACCAAGGGCTCCTCGAGGCGCTTGAAGGGATATCCCGCAGTTGGGTC	307		
Qy	302	CCACCTTTGGACACTGCAGCTGGAGCTGCGGACTTTTGCCACCACTCTGCGACAGAG	361		
Db	308	CCACCTTTGGACACTGCAGCTGGAGCTGCGGACTTTTGCCACCACTCTGCGACAGAG	367		
Qy	362	TGGAAGAACTTGGNAATGGCCCTTGCCCTGCAGGCCACCCAGAGGTGCATCCGGCGCTTCG	421		
Db	368	TGGAAGAACTTGGNAATGGCCCTTGCCCTGCAGGCCACCCAGAGGTGCATCCGGCGCTTCG	427		
Qy	422	CCTCTGCTTTCCAGCGCGGGCAGGAGGGTCTAGTTGCTCTCCCATCTGCAGAGCTTCC	481		
Db	428	CCTCTGCTTTCCAGCGCGGGCAGGAGGGTCTAGTTGCTAGCCATCTGCAGAGCTTCC	487		
Qy	482	TGGAGGTGTCGTACCGCGTTCTACGCCACCTTGCACGACC	522		
Db	488	TGGAGGTGTCGTACCGCGTTCTACGCCACCTTGCACGACC	528		

```

RESULT 13
US-08-469-318-75
; Sequence 75, Application US/08469318
; Patent No. 6022535
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 196
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,318
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/446,872
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 75:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 921 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-469-318-75

```

Query Match 90.98; Score 474.6; DB 3; Length 921;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2002, 08:59:09 ; Search time 262.13 seconds
(without alignments)
3419.026 Million cell updates/sec

Title: US-09-680-514-4_COPY_526_1047
Perfect score: 522
Sequence: 1 GCGCCAAACATATCGCGCTC.....TACGCCACCTGCCCCAGCCC 522

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	522	100.0	1047	17	AAT41786
2	522	100.0	1095	17	AAT41788
3	520.4	99.7	1083	17	AAT41787
4	514	98.5	525	11	AA004482
5	510.8	97.9	525	11	AA004484
6	506	96.9	525	11	AA004481
7	486.8	93.3	525	9	AAN80947
8	486.8	93.3	531	22	AAD19772
9	486.8	93.3	615	22	AAD19771
10	522	100.0	1047	17	AAT41786
11	522	100.0	1095	17	AAT41788
12	520.4	99.7	1083	17	AAT41787
13	514	98.5	525	11	AA004482
14	510.8	97.9	525	11	AA004484
15	506	96.9	525	11	AA004481
16	486.8	93.3	525	9	AAN80947
17	486.8	93.3	531	22	AAD19772
18	486.8	93.3	615	22	AAD19771

10	486.8	93.3	1520	8	AAN70223	Plasmid pBRV2 inse
11	486.8	93.3	1520	8	AAN71320	Sequence encoding
12	486.8	93.3	1521	9	AAN60937	Plasmid pBRV2 inse
13	486.8	93.3	1525	9	AAN81478	Sequence encoding
14	486.8	93.3	1525	10	AAN91086	Plasmid pP12 contg
15	485.2	93.0	1415	8	AAN71089	Sequence of human
16	484.2	92.8	644	22	AAI71848	Recombinant human
17	474.6	90.9	546	16	AAQ97206	pMON13499 DNA sequ
18	474.6	90.9	546	16	AAQ97205	pMON13010 DNA sequ
19	474.6	90.9	546	21	AAQ37772	Human G-CSF mutant
20	474.6	90.9	546	21	AAQ37773	Human G-CSF mutant
21	474.6	90.9	921	16	AAQ97186	pMON13026 DNA enco
22	474.6	90.9	921	16	AAQ97189	pMON13043 DNA enco
23	474.6	90.9	921	16	AAQ97192	pMON13151 DNA enco
24	474.6	90.9	921	16	AAQ97198	pMON13066 DNA enco
25	474.6	90.9	921	21	AAQ3740	Human interleukin-
26	474.6	90.9	921	21	AAQ3743	Human interleukin-
27	474.6	90.9	921	21	AAQ3746	Human interleukin-
28	474.6	90.9	921	21	AAQ3752	Human interleukin-
29	474.6	90.9	923	20	AA238850	Human granulocytic
30	474.6	90.9	966	16	AAQ97184	pMON13058 DNA enco
31	474.6	90.9	966	16	AAQ97185	pMON13060 DNA enco
32	474.6	90.9	966	16	AAQ97187	pMON13063 DNA enco
33	474.6	90.9	966	16	AAQ97191	pMON13045 DNA enco
34	474.6	90.9	966	16	AAQ97193	pMON13152 DNA enco
35	474.6	90.9	966	16	AAQ97182	pMON13034 DNA enco
36	474.6	90.9	966	21	AAQ3736	Human interleukin-
37	474.6	90.9	966	21	AAQ3738	Human interleukin-
38	474.6	90.9	966	21	AAQ3739	Human interleukin-
39	474.6	90.9	966	21	AAQ3741	Human interleukin-
40	474.6	90.9	966	21	AAQ3745	Human interleukin-
41	474.6	90.9	966	21	AAQ3747	Human interleukin-
42	474.6	90.9	1017	16	AAQ97181	pMON15937 DNA enco
43	474.6	90.9	1017	21	AAQ3735	Human interleukin-
44	474.6	90.9	1047	16	AAQ97188	pMON13084 DNA enco
45	474.6	90.9	1047	16	AAQ97190	pMON13044 DNA enco

ALIGNMENTS

RESULT 1	
AAT41786	
ID	AAT41786 standard; DNA; 1047 BP.
XX	AC
XX	AAT41786;
XX	01-JUL-1997 (first entry)
DT	Fusion peptide #1 having G-CSF and TPO activity.
DE	
DE	Fusion protein; human granulocyte colony stimulating factor; hg-CSF;
KW	thrombopoietin; TPO; spacer peptide; blood platelet production;
KW	leukocyte production; anaemia; ds.
XX	
OS	Homo sapiens.
XX	
EH	Key
FT	sig_peptide
FT	1..63
FT	/*tag= a
FT	64..1047
FT	/*tag= b
XX	
PN	WO9634016-A1.
XX	
PD	31-OCT-1996.
XX	
PF	26-APR-1996;
XX	96WO-JP01157.
XX	
PR	26-APR-1995;
XX	95JP-0102625.
PA	(KYOW) KYOWA HAKKO KOGYO KK.
XX	

PI Anazawa H, Konishi N, Shiotsu Y, Tamaoki T, Terasaki Y;
PI Uchida K, Yamasaki M, Yamashita K, Yokoi H;
XX
XX WPI: 1996-497573/49.
DR P-PSDB; AAW00377.
XX
XX Fusion peptide having G-CSF and with thrombopoietin activity -
PT optionally chemically modified with a poly:alkylene glycol, used for
PT treatment of anaemia
XX
XX Claim 3; Page 46-48; 71pp; Japanese.
XX
XX The sequences given in AAT41786-88 encode fusion proteins which are
CC composed of a peptide having human granulocyte colony stimulating
CC factor (hG-CSF) activity fused with a peptide having thrombopoietin
CC (TPO) activity, opt. via a spacer peptide. Peptides derived from
CC these by deletion, insertion or substitution of one or more amino
CC acid residues are included within the scope of the invention. The
CC fusion peptides stimulate blood platelet and leukocyte production
CC and are useful in the treatment of anaemia.
XX
XX Sequence 1047 BP; 192 A; 351 C; 288 G; 216 T; 0 other;

Query Match 100.0%; Score 522; DB 17; Length 1047;
Best Local Similarity 100.0%; Pred. No. 2.7e-109; Indels 0; Gaps 0;
Matches 522; Conservative 0; Mismatches 0;

QY 1 GCGCCACATATCGCGCTCGAGTCTACACAGAGTTCCTTTTAAAGCTTAGACAA 60
DB 526 GCGCCACATATCGCGCTCGAGTCTACACAGAGTTCCTTTTAAAGCTTAGACAA 585
QY 61 GTGAGAGATCCAGGGGATGGCGAGCGCTCAGAGAGCTGTGCGCACTACAG 120
DB 586 GTGAGAGATCCAGGGGATGGCGAGCGCTCAGAGAGCTGTGCGCACTACAG 645
QY 121 CTGTGCCACCCGAGGAGCTGGTGTCTGGACACTCTCTGGCATCCCTGGGCTCCC 180
DB 646 CTGTGCCACCCGAGGAGCTGGTGTCTGGACACTCTCTGGCATCCCTGGGCTCCC 705
QY 181 CTGAGCAGCTGCCCGAGCGCCCTCGAGCTGGCAGGCTGCTTGAAGCACTTCCATAGC 240
DB 706 CTGAGCAGCTGCCCGAGCGCCCTCGAGCTGGCAGGCTGCTTGAAGCACTTCCATAGC 765
QY 241 GCGCTTTTCTCTACAGGGGCTCTCGAGCGCTGGAAGGATCTCCCGAGTGGGT 300
DB 766 GCGCTTTTCTCTACAGGGGCTCTCGAGCGCTGGAAGGATCTCCCGAGTGGGT 825
QY 301 CCCACCTTGGACACTGCGAGCTGGAGCTGCCGACTTTGCCACCACTCTGGCAGCAG 360
DB 826 CCCACCTTGGACACTGCGAGCTGGAGCTGCCGACTTTGCCACCACTCTGGCAGCAG 885
QY 361 ATGGAAGAACTGGGAATGGCCCTTGCCTCGACGCCACCCAGGCTGCCATGCCGCTTC 420
DB 886 ATGGAAGAACTGGGAATGGCCCTTGCCTCGACGCCACCCAGGCTGCCATGCCGCTTC 945
QY 421 GCGCTTCTTCCAGCGCGGGAGGAGGCTCTAGTTCCTTCCATCTGCGAGCTTC 480
DB 946 GCGCTTCTTCCAGCGCGGGAGGAGGCTCTAGTTCCTTCCATCTGCGAGCTTC 1005
QY 481 CTGAGAGTGTCTACCGCTTCTACGCGACTTGTGCCAGGCC 522
DB 1006 CTGAGAGTGTCTACCGCTTCTACGCGACTTGTGCCAGGCC 1047

RESULT 2
AAT41788
ID AAT41788 standard; DNA: 1095 BP.
XX
XX AC AAT41788;
XX
XX DT 01-JUL-1997 (first entry)
XX

DE Fusion peptide #3 having G-CSF and TPO activity.
XX
XX Fusion protein; human granulocyte colony stimulating factor; hG-CSF;
KW thrombopoietin; TPO; spacer peptide; blood platelet production;
KW leukocyte production; anaemia; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH sig_peptide 1..63
FT /*tag= a
FT mat_peptide 64..1095
FT /*tag= b
XX
XX W09634016-A1.
XX
XX 31-OCT-1996.
XX
XX 26-APR-1996; 96WO-JP01157.
XX
XX 26-APR-1995; 95JP-0102625.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX
XX Anazawa H, Konishi N, Shiotsu Y, Tamaoki T, Terasaki Y;
PI Uchida K, Yamasaki M, Yamashita K, Yokoi H;
XX
XX WPI: 1996-497573/49.
DR P-PSDB; AAW00379.
XX
XX Fusion peptide having G-CSF and with thrombopoietin activity -
PT optionally chemically modified with a poly:alkylene glycol, used for
PT treatment of anaemia
XX
XX Claim 3; Page 52-54; 71pp; Japanese.
XX
XX The sequences given in AAT41786-88 encode fusion proteins which are
CC composed of a peptide having human granulocyte colony stimulating
CC factor (hG-CSF) activity fused with a peptide having thrombopoietin
CC (TPO) activity, opt. via a spacer peptide. Peptides derived from
CC these by deletion, insertion or substitution of one or more amino
CC acid residues are included within the scope of the invention. The
CC fusion peptides stimulate blood platelet and leukocyte production
CC and are useful in the treatment of anaemia.
XX
XX Sequence 1095 BP; 194 A; 362 C; 311 G; 228 T; 0 other;

Query Match 100.0%; Score 522; DB 17; Length 1095;
Best Local Similarity 100.0%; Pred. No. 2.7e-109; Indels 0; Gaps 0;
Matches 522; Conservative 0; Mismatches 0;

QY 1 GCGCCACATATCGCGCTCGAGTCTACACAGAGTTCCTTTTAAAGCTTAGACAA 60
DB 574 GCGCCACATATCGCGCTCGAGTCTACACAGAGTTCCTTTTAAAGCTTAGACAA 633
QY 61 GTGAGAGATCCAGGGGATGGCGAGCGCTCAGAGAGCTGTGTCACCTACAG 120
DB 634 GTGAGAGATCCAGGGGATGGCGAGCGCTCAGAGAGCTGTGTCACCTACAG 693
QY 121 CTGTGCCACCCGAGGAGCTGGTGTCTGGACACTCTCTGGCATCCCTGGGCTCCC 180
DB 694 CTGTGCCACCCGAGGAGCTGGTGTCTGGACACTCTCTGGCATCCCTGGGCTCCC 753
QY 181 CTGAGCAGTGCCTCCAGCGCCCTCGAGCTGGCAGGCTGCTTGAAGCACTTCCATAGC 240
DB 754 CTGAGCAGTGCCTCCAGCGCCCTCGAGCTGGCAGGCTGCTTGAAGCACTTCCATAGC 813
QY 241 GCGCTTTTCTCTACAGGGGCTCTCGAGCGCTGGAAGGATCTCCCGAGTGGGT 300
DB 814 GCGCTTTTCTCTACAGGGGCTCTCGAGCGCTGGAAGGATCTCCCGAGTGGGT 873
QY 301 CCCACCTTGGACACTGCGAGCTGGAGCTGCCGACTTTGCCACCACTCTGGCAGCAG 360

|||||
Db 874 cccacctggacacactgcagctgacgtgcgcgactttgccaccacatctggcagcag 933
Qy 361 ATGGAAGAACTGGGAATGCGCCCTGCGCTGCAGAGCCACCCAGGGTGCATGCCGCGCTTC 420
Db 934 atggaagaacg99gaatggccctgcctgcagccacccaggggtgcacatgcggcgcttc 993
Qy 421 GCCTCTGCTTTCCAGCGCGGCGAGAGGGGTCTCTAGTTGGCTTCCCATCTGCAGAGCTTC 480
Db 994 gccctgctttccagcgccggcagagggtcctagttgctcccatctgcagagcttc 1053
Qy 481 CTGAGAGTGTCTACCGCGTTCCTAGCCACCTTGGCCAGCC 522
Db 1054 ctggagggtgctgacccgcttctacgccacacttgccagccc 1095

RESULT 3
AA041787
ID AA041787 standard; DNA; 1083 BP.
XX
AC AA041787;
DT 01-JUL-1997 (first entry)
XX
DE Fusion peptide #2 having G-CSF and TPO activity.
XX
KW Fusion protein; human granulocyte colony stimulating factor; hg-CSF;
KW thrombopoietin; TPO; spacer peptide; blood platelet production;
KW leukocyte production; anaemia; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT sig_peptide 1..63 /*tag= a
FT mat_peptide 64..1083 /*tag= b
FT
FN W09634016-A1.
XX
XX 31-OCT-1996.
XX
PF 26-APR-1996; 96WO-JP01157.
XX
PR 26-APR-1995; 95JP-0102625.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Anazawa H, Konishi N, Shiotsu Y, Tamaoki T, Terasaki Y;
PI Uchida K, Yamasaki M, Yamashita K, Yokoi H;
XX
WPI: 1996-497573/49.
DR P-PSDB; AAW00378.
XX

Fusion peptide having G-CSF and with thrombopoietin activity -
PT optionally chemically modified with a poly(alkylene glycol, used for
PT treatment of anaemia
XX
PS Claim 3; Page 49-51; 71pp; Japanese.
XX
CC The sequences given in AA041786-88 encode fusion proteins which are
CC composed of a peptide having human granulocyte colony stimulating
CC factor (hg-CSF) activity fused with a peptide having thrombopoietin
CC (TPO) activity, opt. via a spacer peptide. Peptides derived from
CC these by deletion, insertion or substitution of one or more amino
CC acid residues are included within the scope of the invention. The
CC fusion peptides stimulate blood platelet and leukocyte production
CC and are useful in the treatment of anaemia.
XX
SQ Sequence 1083 BP; 197 A; 355 C; 305 G; 226 T; 0 other;

Query Match 99.7%; Score 520.4; DB 17; Length 1083;

Best Local Similarity 99.8%; Pred. No. 6.3e-109;
Matches 521; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GCGCCAAATATCGCGCCTCGAGTCTACACAGAGCTTCCCTTTTAAAAAGCTTAGAGCAA 60
Db 562 gcaccaacatcgcgcctcagctctacacagagctccttttaaaagcttagagcaa 621
Qy 61 CTGAGAAAGATCCAGGCGGATGGCGCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAA 120
Db 622 gtgaggaagatccaggcgatggcgagcgctcccaaggagaagctgtgccaacctacaag 681
Qy 121 CTGTGCCACCCCGAGAGAGCTGGTGTCTGCTCGGACACTCTCTGGGCAATCCCTGGGCTCC 180
Db 682 ctgtgccaccccgaggagctggtgctgctcggaacctctctgggcatccctgggctccc 741
Qy 181 CTGAGCAGCTGCCCGCAGCCAGCCCTGCAGCTGGCAGCGCTGCTTGCAGCCAACTCCATAGC 240
Db 742 ctgagcagctgccccagccagccgctgcagctggcaggtgctgagccaaactccatagc 801
Qy 241 GGCCTTTTCTCTACAGGGGCTCCTGCAGGCGCTTGAAGGATCTCCCGAGTTGGGT 300
Db 802 ggccttttctctaccagggtcctcctgcagccctggaaggatctccccgagttgggt 861
Qy 301 CCCACCTTGGACACACTGCAGCTGCAGCTGCCGCGACTTGGCCACCATCTGGCAGCAG 360
Db 862 cccaccttggacacactgcagctggcagctgcgcgactttgccaccacatctggcagcag 921
Qy 361 ATGGAAGAACTGGGAATGCGCCCTGCGCTGCAGAGCCACCCAGGGTGCATGCCGCGCTTC 420
Db 922 atggaagaactgggaatggccctgcctgcagccacccaggggtgccaatgcggcgcttc 981
Qy 421 GCCTCTGCTTTCCAGCGCGGCGAGAGGGGTCTCTAGTTGGCTTCCCATCTGCAGAGCTTC 480
Db 982 gccctgctttccagcgccggcagagggtcctagttgctcccatctgcagagcttc 1041
Qy 481 CTGAGAGTGTCTACCGCGTTCCTAGCCACCTTGGCCAGCC 522
Db 1042 ctggagggtgctgacccgcttctacgccacacttgccagccc 1083

RESULT 4
AA04482
ID AA04482 standard; DNA; 525 BP.
XX
AC AA04482;
DT 04-OCT-1990 (first entry)
XX
DE Plasmid pASN6 encoding hg-CSF[ND28N6].
XX
KW Granulocyte stimulating factor; O-linked glycosylation; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_difference 1..3 /*tag= a
FT /*label=synthetic_mutation
FT /*note="old seq (acc)"
FT misc_difference 7..18 /*tag= b
FT /*label=synthetic_mutation
FT /*note="old seq (ctgggcccctgcc)"
FT misc_difference 49..51 /*tag= c
FT /*label=synthetic_mutation
FT /*note="old seq (tgc)"
XX
PN EP370205-A.
XX
PD 30-MAY-1990.
XX
PF 28-SEP-1989; 89EP-0117981.

```
XX 29-SEP-1989; 89JP-0245705.
XX (KYOW ) KYOWA HAKKA KK.
XX Sasaki K, Nishi T, Yasumuru S, Sato M, Itoh S;
XX WPI; 1990-165029/22.
XX P-PSDB; AAR05114.
XX Polypeptide(s) with added carbohydrate chains - formed by
PT modification of amino acid sequence, used to improve
PT physiochemical properties and/or activities.
XX PS Disclosure; ; 30pp; English.
XX The sequence encodes a deriv. of mature hG-CSF, designated
CC hG-CSF[ND28N] which has 6 amino acid substns. This results in
CC one extra O-linked glycosylation site and one new N-linked
CC glycosylation site giving the new protein improved stability and
CC activity.
CC See also AAQ04481-86.
XX SQ Sequence 525 BP; 94 A; 183 C; 148 G; 100 T; 0 other;

Query Match 98.5%; Score 514; DB 11; Length 525;
Best Local Similarity 99.0%; Pred. No. 1.6e-107;
Matches 517; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCGCCACATATCGCGCTCGAGTCTACACAGAGCTTCCTTTTAAAGCTTAGACAA 60
DB 1 gcaactcgtatcgcaactcgagctgtaccacagagcttcttttaaaagcttagagcaa 60

QY 61 CTGAGGAGATCCAGGCGATGGCGAGCTCCAGGAGAGCTGTGCCACCTACAG 120
DB 61 gtgagggaatcccaaggcgatgcgacgagctccagagagctgtgtgccacctacaag 120

QY 121 CTGTGCCACCCGAGGAGCTGTGCTCTCGGACACTCTCTGGGCATCCCTGGGCTCCC 180
DB 121 ctgtgccacccgagagctgtgtgctctcgagacactctctgggcatccctgggctccc 180

QY 181 CTGAGAGCTGCCCGACGAGCGCTCGAGCTGCGAGGCTGTGAGCCAACTCCATAGC 240
DB 181 ctgagcagctgcccgacgagcgccctgagctgagcagctgttgagccaaactccatagc 240

QY 241 GGCCTTTTCTTCTACAGGGGCTCTCGAGCGCTCGAGGGATCTCCCGAGTTGGGT 300
DB 241 ggccttttcttctacaggggctctcgagcgctcgaggggagatctccccccgagttgggt 300

QY 301 CCCACCTGGACACACTGCAGCTGGAGCTGCGCGACTTTTGCCACCACTCTGGCAGCAG 360
DB 301 cccacctggacacactgcagctggagctgagcgtcgccgacttgccaccacctctggcagag 360

QY 361 ATGAAGAAGCTGGGAATGGCCCTGCGCTCGAGCCACCCAGGGTGCATGCGCGGCTTC 420
DB 361 atgaagaagctgggaatggccctcgctcgagccacccaggggtgccatgcccgccttc 420

QY 421 GCCTCTGCTTCCAGCGCGGAGGAGGGTCTTAGTGCTCCCTCCATCTGCGAGAGCTTC 480
DB 421 gcctctgcttccagcgccgagggaggggtcttagtgctccctccatctgagagagcttc 480

QY 481 CTGAGGTGTGTACCGGCTTCTACGCCACCTTGCCAGGCC 522
DB 481 ctgagggtgtgtaccgcttctacgccacaccttgccagccc 522

RESULT 5
AAQ04484
ID AAQ04484 standard; DNA; 525 BP.
XX AAQ04484;
XX AAQ04484;
```

```
DT 04-OCT-1990 (first entry)
XX Plasmid pASN145 encoding hG-CSF[ND28N145].
XX KW Granulocyte stimulating factor; glycosylation; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT misc_difference 1..3 /*tag= a
FT /label=synthetic_mutation
FT /note="old seq (acc)"
FT misc_difference 7..18 /*tag= b
FT /label=synthetic_mutation
FT /note="old seq (ctgggcccctgccc)"
FT misc_difference 49..51 /*tag= c
FT /label=synthetic_mutation
FT /note="old seq (tgc)"
FT misc_difference 433..435 /*tag= d
FT /label=synthetic_mutation
FT /note="old seq (cag)"
FT misc_difference 439..441 /*tag= e
FT /label=synthetic_mutation
FT /note="old seq (cgg)"
XX EP370205-A.
XX 30-MAY-1990.
XX 28-SEP-1989; 89EP-0117981.
XX 29-SEP-1989; 89JP-0245705.
XX (KYOW ) KYOWA HAKKA KK.
XX Sasaki K, Nishi T, Yasumuru S, Sato M, Itoh S;
XX WPI; 1990-165029/22.
XX P-PSDB; AAR05115.
XX Polypeptide(s) with added carbohydrate chains - formed by
PT modification of amino acid sequence, used to improve
PT physiochemical properties and/or activities.
XX PS Disclosure; ; 0pp; English.
XX The sequence encodes a deriv. of mature hG-CSF, designated
CC hG-CSF[ND28N145] which has 7 amino acid substns. This results in
CC one extra O-linked glycosylation site and two new N-linked
CC glycosylation site giving the new protein improved stability and
CC activity.
CC See also AAQ04481-86.
XX SQ Sequence 525 BP; 96 A; 183 C; 145 G; 101 T; 0 other;

Query Match 97.9%; Score 510.8; DB 11; Length 525;
Best Local Similarity 98.7%; Pred. No. 8.6e-107;
Matches 515; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 GCGCCACATATCGCGCTCGAGTCTACACAGAGCTTCCTTTTAAAGCTTAGACAA 60
DB 1 gcaaccaacatatcgcgctcgagctgtaccacagagcttcttttaaaagcttagagcaa 60

QY 61 GTGAGGAGATCCAGGCGATGGCGAGGCTCCAGGAGAGCTGTGTGCCACCTACAG 120
DB 61 gtgagggaagatccaggcgatgcgagcgctccagagagagctgtgtgccacctacaag 120
```

```
QY 121 CTGTGCCACCCGAGGAGCTGTGTGCTGCTGACACTCTCTGGGATCCCTGGGCTCCC 180
Db 121 cttgtgccacccgaggagctgtgtgtctgcgcgacactctctggcctccctgggtctccc 180
QY 181 CTGACAGCTGCCCGAGCCCTGCGAGCTGGCAGGCTGCTTGAGCCAACTCCATAGC 240
Db 181 ctgtgcagctgcccgagccctgacgctggcagctgtgttgagccaaactccatagc 240
QY 241 GGCCTTTTCTCTACAGGGCTCTGTGAGGCTCTGGAAGGCTCTCCCGGAGTTGGGT 300
Db 241 ggccttttctctaccagggtctctgcagccttggaaggatctccccagattgggt 300
QY 301 CCCACTTGGACACACTGTCAGCTGACGTCGCGGCTTTGCCACCACTATCTGGCAGCAG 360
Db 301 cccacttggacacactgctcagctgacgtcgcgactttggcaccaccatctggcagcag 360
QY 361 ATGGAGAACTGGGAATGCCCTGCTCCCTGTCAGCCACCCAGGCTGCCATGCCGCTTC 420
Db 361 atggaagaactgggaatggccctgctccctgcagccacccaggtgccaatgcccgccttc 420
QY 421 GCCTCTGCTTTCCAGCGCGGCGAGGGGTCTCTAGTTGCTTCCCATCTGCAGAGCTTC 480
Db 421 gcctctgtcttcaatcgatcgagcagagggttccctagttgctcccatctgcagagcttc 480
QY 481 CTGGAGGTGTCTACCGGCTTCTACGCCACCTTGCCAGCC 522
Db 481 ctggagggtgtctaccggttctacgccacacttgcccagccc 522
```

RESULT 6

```
ID AAQ04481
AC AAQ04481;
XX 04-OCT-1990 (first entry)
XX Plasmid pAS28 encoding hg-CSF[ND28].
DE Granulocyte stimulating factor; O-linked glycosylation; ss.
KW Homo sapiens.
OS
```

```
Key Location/Qualifiers
misc_difference 1..3
FT /*tag= a
FT /label=synthetic_mutation
FT /note="old seq (acc)"
misc_difference 7..15
FT /*tag= b
FT /label=synthetic_mutation
FT /note="old seq (ctgggacct)"
misc_difference 49..51
FT /*tag= c
FT /label=synthetic_mutation
FT /note="old seq (tgc)"
```

```
PN EP370205-A.
XX 30-MAY-1990.
XX 28-SEP-1989; 89EP-0117981.
XX 29-SEP-1989; 89JP-0245705.
XX (KYOW ) KYOWA HAKKA KK.
XX Sasaki K, Nishi T, Yasumuru S, Sato M, Itoh S;
XX WPI: 1990-165029/22.
XX P-PSDB; AAR05113.
XX Polypeptide(s) with added carbohydrate chains - formed by
```

```
PT modification of amino acid sequence, used to improve
PT physiochemical properties and/or activities.
PS Disclosure; ; 30pp; English.
XX The sequence encodes a deriv. of mature hg-CSF, designated
CC hg-CSF[N28] which has 5 amino acid substns. This results in
CC one extra O-linked glycosylation site giving the new protein
CC improved stability and activity.
CC See also AAQ04482-86.
XX Sequence 525 BP; 94 A; 184 C; 149 G; 98 T; 0 other;
```

```
Query Match 96.9%; Score 506; DB 11; Length 525;
Best Local Similarity 98.1%; Pred. No. 1..le-105;
Matches 512; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 1 GCGCCAAACATATCGGCGCTCGAGTCTACACAGAGCTCTCTTTTAAAGCTTAGAGCAA 60
Db 1 gcaccaacatatcgcgctcgcagctctaccacagagcttcttttaaaaagcttagagcaa 60
QY 61 GTGAGGAAGATCCAGGGCGATGGCGAGCGCTCCAGGAGAAAGCTGTGCTGCCACTACAG 120
Db 61 gtgaggaagatccaggcgatggcgagcgtccaggagagagctcgtgccacctacaag 120
QY 121 CTGTGCCACCCGAGGAGCTGTGTGCTGCTGACACTCTCTGGGATCCCTGGGCTCCC 180
Db 121 cttgtgccaccccgaggagctgtgtgtcttcggacactctctgggcatcccttgggctccc 180
QY 181 CTGAGCAGTGTGCCCGAGCCCTGTCAGCTGGCAGGCTGCTTGAGCCAACTCCATAGC 240
Db 181 ctgagcagctgcccgagccctgcagctggcagcgtctgtgagccaaactccatagc 240
QY 241 GGCCTTTTCTCTACAGGGCTCTCTGAGGCTCTGAGGCTCTGGAAGGCTCTCCCGGAGTTGGGT 300
Db 241 ggccttttctctaccagggtctctgcagccttggaaggatctccccagattgggt 300
QY 301 CCACCTTGGACACACTGTCAGCTGTCGCGGCTTTGCCACCACTATCTGGCAGCAG 360
Db 301 cccacttggacacactgctcagctgacgtcgcgactttggcaccaccatctggcagcag 360
QY 361 ATGGAAGAACTGGGAATGCCCTGCTCCCTGTCAGCCACCCAGGCTGCCATGCCGCTTC 420
Db 361 atggaagaactgggaatggccctgctccctgcagccacccaggtgccaatgcccgccttc 420
QY 421 GCCTCTGCTTTCCAGCGCGGCGAGGGGTCTCTAGTTGCTTCCCATCTGCAGAGCTTC 480
Db 421 gcctctgtcttcaatcgagcagagggttccctagttgctcccatctgcagagcttc 480
QY 481 CTGGAGGTGTCTACCGGCTTCTACGCCACCTTGCCAGCC 522
Db 481 ctggagggtgtctaccggttctacgccacacttgcccagccc 522
```

RESULT 7

```
AA080947
ID AA080947 standard; DNA; 525 BP.
XX AC AA080947;
XX 15-NOV-1990 (first entry)
XX G-CSF gene isolated from peripheral blood macrophages.
XX Granulocyte colony stimulating factor; hg-CSF; macrophage;
XX tumour therapy; leukaemia; ss.
XX Homo sapiens.
XX EP272703-A.
XX 29-JUN-1988.
PD
```


|||||
Db 148 ctgctcgacactctctgggcatccctcctgggtccctcctgagcagctgccccagcagcc 207
QY 205 CTGACAGTGGCAGGCTGTGAGCAACTCCATAGCGGCTTTCTCTACCAAGGGGCTC 264
Db 208 ctgagctggcaggtgttgagcaactccatagcgccctttctctaccaggggtc 267
QY 265 CTGACAGGCTTGGAGGATCTCCCGAGTTGGGTGCCACCTTGGACACACTGCAGCTG 324
Db 268 ctgcaaggccttgaaggatctccccagattgggtccacccttggacacactgcagctg 327
QY 325 GACGTGCGCGACTTTGCCACCACTCTGCGAGCAGATGGAGAACTGGGAATGCCCTC 384
Db 328 gacgtcgccgactttgcccaccactctggcagcagatggaagaactggaatggccct 387
QY 395 GCCCTGACGCCACCCAGGCTGCCATGCGGCTTCGCTCTGCTTCCAGGCCGGGCA 444
Db 398 gccctgcaagccaccaggtgccaatgcgcgcttcgctctgcttccagcgccgga 447
QY 445 GGAGGGTCTTAGTTGGCTCTCCCATCTGCAGAGCTTCCTGGAGGTGCTGCTACCGCTTCTA 504
Db 448 ggaagggtcctagtgtcctccatctgcagagcttcttggaggtgtctaccgcgttcta 507
QY 505 CGCCACCTTGGCCAGCC 522
Db 508 cgcacacttggccagcc 525

RESULT 9

AAD19771
ID AAD19771 standard; DNA; 615 BP.
XX
AC
XX
AC AAD19771;
DT 18-DEC-2001 (first entry)
XX
XX Human hg-CSF gene inserted into plasmid p19CSFm.
DE
KW Human; granulocyte colony stimulating factor; hg-CSF; protease;
KW Factor Xa; kanamycin resistance; endoxylanase signal peptide; ds.
XX
XX Homo sapiens.

Key	Location/Qualifiers
CDs	88..612
FT	/*tag= a
FT	/product= "Human hg-CSF protein"
FT	/note= "CDS does not include start codon"
FT	/partial

XX
XX WO200173081-A1.
XX
XX 04-OCT-2001.
XX
XX 31-MAR-2001; 2001WO-KR00549.
XX
XX 31-MAR-2000; 2000KR-0017052.
XX
XX (ROAD) KOREA ADV INST SCI & TECHNOLOGY.

XX
XX Lee S, Jeong K;
XX
XX WPI; 2001-616523/71.
XX
XX P-PSDB; AAE12153.

XX
XX Recombinant plasmid vector comprising an endoxylanase signal sequence,
XX human granulocyte colony stimulating factor gene and other components,
XX when transformed into microorganism useful for preparing the colony
XX stimulating factor -

XX
XX Example 1; Fig 3; 50pp; English.

XX
XX The invention relates to an Escherichia coli producing and secreting

CC human granulocyte colony stimulating factor (hg-CSF), more specifically,
CC to a recombinant plasmid constructed to express secretory hg-CSF in
CC E. coli, an E. coli transformed with that plasmid to secrete hg-CSF,
CC and a process for preparing hg-CSF using the transformed hg-CSF. The
CC recombinant plasmid vector comprises of a kanamycin resistance gene, a
CC promoter, an endoxylanase signal sequence, a nucleotide sequence coding
CC for an oligopeptide consisting of 13 amino acids including 6 consecutive
CC histidine residues and a hg-CSF. E.coli transformed with recombinant
CC plasmid vector is useful for preparing hg-CSF. The method comprises
CC culturing the microorganism to obtain a hg-CSF fusion protein and
CC treating the fusion protein with a protease preferably Factor Xa, to
CC obtain a hg-CSF, where the fusion protein is obtained from the culture
CC by employing Ni-column. The present sequence is human hg-CSF gene
CC inserted into plasmid p19CSFm.

XX
SQ Sequence 615 BP; 105 A; 219 C; 178 G; 113 T; 0 other;

Query Match 93.3%; Score 486.8; DB 22; Length 615;
Best Local Similarity 98.6%; Pred. No. 2.4e-101;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 25 CTACCACAGAGCTTCCTTTTAAAAAGCTTAGAGCAAGTAGAGAAATCCAGGGCGATGGC 84
Db 112 ctgccccagagcttctcctcaagtgtcttagagcaagtgaggaagatccaggcgatgac 171
QY 85 GCAGCGCTCCAGGAGAGCTGTGCGCACTACAGCTGCCACCCGAGGAGCTGGTG 144
Db 172 gcagcgctccaggaagctgtgtccacctacaagctgtgccaccctgagagctggtg 231
QY 145 CTGCTCGGACACTCTCTGGGATCCCTGGGCTCCCTCGAGCAGCTGCCAGCCAGGCC 204
Db 232 ctgctcgacactctctggcatccctggctccctcctgagcagctgccccagcagcc 291
QY 205 CTGACAGTGGCAGGCTGCTTGGACCAACTCCATACCGGCTTTTCTCTACCAAGGGGCTC 264
Db 292 ctgcagctggcaggtgtgtgagcaactccatagcggcctttctctaccaggggtc 351
QY 265 CTGACAGGCTTGGAGGATCTCCCGAGTTGGGTCCACCTTGGACACACTGCAGCTG 324
Db 352 ctgcagggcccttgaaggatctccccagattgggtccacccttggacacactgcagctg 411
QY 325 GACGTGCGCGACTTTGGCCACCACCATCTGGCAGCAGATGGAGAACTGGGAATGCCCTC 384
Db 412 gacgtcgccgactttgcccaccactctgagcagatggaagaactgggaatggccct 471
QY 395 GCCCTGACAGCCACCCAGGCTGCCATGCGGCTTCGCTCTGCTTCCAGGGCGGCGCA 444
Db 472 gccctgcagccaccaggtgccaatgcgcgcttcgctctgcttccagcgccggca 531
QY 445 GGAGGGTCTTAGTTGGCTTCCTCCATCTGCAGAGCTTCCTGGAGGTGCTGCTACCGCTTCTA 504
Db 532 ggaagggttccctagttgctcccatctgcagagcttccctggagggtgtctaccgcgttcta 591
QY 505 CGCCACCTTGGCCAGGCC 522
Db 592 cgcacacttggccagcc 609

RESULT 10

AAN70223
ID AAN70223 standard; DNA; 1520 BP.
XX
XX
AC AAN70223;

XX
XX 25-APR-1991 (first entry)
XX
XX Plasmid pBRV2 insert.

XX
XX Human granulocyte colony stimulating factor; G-CSF; leukaemia; ss.
XX
XX Homo sapiens.

Key Location/Qualifiers
CDS 31..645
/tag= a

EP220520-A.

06-MAY-1987.

30-SEP-1986; 86EP-0113446.

30-SEP-1985; 85JP-0217150.

17-JUL-1986; 86JP-0166710.

17-JUL-1986; 86JP-0166709.

(CHUS) CHUGAI SEIYAKU KK.

Yamazaki T, Nagata S, Tsuchiya M;

WPI; 1987-124182/18.

P-PSDB; AAP70162.

Polypeptide with human granulocyte colony stimulating factor activity - is obtd. by cultivating transformant formed by recombinant DNA procedures.

Disclosure; Fig 4; 73pp; English.

The plasmid was isolated from a cDNA library prepd. from CHU-2 cells, a human oral cavity tumour cell line, using a 1500 bp insert from pBRG4 (AAN70222) and probe LC (AAN71351). The plasmid was used to prepare recombinant expression plasmids for the prodn. of h G-CSF.

See also AAN70221-N70224 and AAN71349-N71351.

Sequence 1520 BP; 303 A; 487 C; 402 G; 328 T; 0 other;

Query Match 93.3%; Score 486.8; DB 8; Length 1520;
Best Local Similarity 98.6%; Pred. No. 2.7e-101;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

25 CTACACAGAGCTTCCCTTTTAAAGCTTAGAGAGTCCAGGCGGATGGC 84
145 ctgccccagagcttcctgcaagcttagagcaagtgaggagatccaggcgatggc 204

85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAAGCTGTGCCACCCGAGGAGCTGGTG 144
205 gcagcgctccagagagctgtgtgcccacctacaagctgtgccaccgccgagagctggg 264

145 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCCGAGCGGCC 204
265 ctgctcgagacctctctggcatccctggctccctgagcagctgccccagcgcc 324

205 CTGAGCTGGAGGCTCTTGGAGCACTCCATAGAGGCTTCTTCTACAGGGGCTC 264
325 ctgagctggcaggtgcttgagcgaactccatagcgcccttctctaccaggggctc 384

265 CTGAGGCGCTTGGAGGAGTCTCCCGAGTGTGGTCCACCTTGACACACTGCAGCTG 324
385 ctgagcgcccttgagggagatctcccccagttggtgtccacacctggacacactgcagctg 444

325 GACCTCGCGCACTTTGCCACCACTCTGCGACAGATGGAAGTGGGAATGGCCCT 384
445 gacgtccggaacttggccaccaccttgccagcagatggaagaactgggaatggccct 504

395 GCGCTGAGCGCCACCGAGGTGCATGCGGCTTCCCTCTGCTTCCAGCGCCGGCA 444
505 gccctgagcgccaccaggtgcatgtcgccgttcgctctgcttccagcgcgga 564

445 GGAGGGTCTAGTGTCTCCATCTGACAGCTTCTGAGAGTGTCTGACGCTTCTA 504
565 ggaaggggtctagttgcttccatctgacagagcttctgaggtgtcggtaccgcttcta 624

QY 505 CGCCACCTTGCCCGAGGCC 522
Db 625 cgccaccttgcccgagccc 642

RESULT 11
AAN71320
ID AAN71320 standard; cDNA; 1520 BP.
XX.
AC AAN71320;
XX
DT 23-APR-1991 (first entry)
XX
DE Sequence encoding human granulocyte colony stimulating factor (hGCSF) in pBRV2.
XX
KW Leukopenia therapy; neutropenia; eosinopenia; lymphopenia; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 31..120
FT /tag= a
FT mat_peptide 121..645
FT /tag= b

EP217404-A.
XX
PD 08-APR-1987.
XX
PF 03-OCT-1986; 86EP-0113671.
XX
PR 02-JUN-1986; 86JP-0125660.
PR 04-OCT-1985; 85JP-0220450.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
XX
PI Tamura M, Nomura H, Hattori K, Ono M;
XX
WPI; 1987-095431/14.
DR P-PSDB; AAP71384.
XX
PT Leukopenia treating agent esp. for producing fully mature
PT neutrophils - contains human granulocyte colony stimulating
PT factor obtd. by recombinant DNA methods etc.
XX
PS Claim 7; Fig 2; 34pp; English.
XX
CC The hGCSF has the following properties: (i) mol. wgt. 19000 +/- 1000
CC (by SDS-PAGE); (ii) isoelectric pt. at 5.5, 5.8 or 6.1 each +/- 0.1;
CC (iii) UV max. absorption at 280nm and min. at 250nm. Prodn. of the
CC gene, vector etc. is described in JP 269455, 269456, 270838 and
CC 270839, each of 1985.
XX
SQ Sequence 1520 BP; 305 A; 488 C; 401 G; 326 T; 0 other;

Query Match 93.3%; Score 486.8; DB 8; Length 1520;
Best Local Similarity 98.6%; Pred. No. 2.7e-101;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

25 CTACACAGAGCTTCCCTTTTAAAGCTTAGAGAGTCCAGGCGGATGGC 84
145 ctgccccagagcttcctgcaagcttagagcaagtgaggagatccaggcgatggc 204

85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAAGCTGTGCCACCCGAGGAGCTGGTG 144
205 gcagcgctccagagagctgtgtgcccacctacaagctgtgccaccgccgagagctgg 264

145 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCCGAGCGGCC 204
265 ctgctcgagacctctctgagagcttctgaggtgtccctgagcagctgccccagcgcc 324

Qy 205 CTGACGCTGGCAGGCTGCTTGGAGCAACTCCATAGCGCCCTTTCTCTACACAGGGGCTC 264
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
375 ctgcagctggcagcgtgcttagcacaactccatagcgccttttctctaccaggggctc 384
Qy 265 CTGACAGCCCTGGAGAGGATCTCCCGAGTTGGTCCACTTGGACACACTGCAGCTG 324
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
385 ctgcagcgcctggaaggatctcccccgagttgggtccacacttggacacactgagctg 444
Qy 325 GACCTGCGCGACTTGGCCACCACTCTGGCAGCAGATGGAGAACTGGGAATGCCCCCT 384
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
445 gacgtcgccgacttggccaccacatctggcagcagatgggaactgggaatggccct 504
Qy 385 GCCCTGACGACCCACCGAGGTGCCATCGCGGCTTGCCTCTGCTTCCAGCGCGGGCA 444
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
505 gccctgcagccaccacaggtggccatgccgcttgcctctgtttccagcgccggca 564
Qy 445 GGAGGGTCTAGTTGCTCCACTCTGCAGAGCTTCTGGAGGTCTGTACCGGTTCTTA 504
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
565 ggaggggtcctagttgctctccatctgcagagcttctctggaggtgtcgtaccggttcta 624
Qy 505 CGCCACCTTGGCCAGCCC 522
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
625 cgcacacttggccagccc 642

RESULT 12

AAAN60937
ID AAAN60937 standard; DNA; 1521 BP.
XX
AC AAAN60937;
XX
DT 18-OCT-1991 (first entry)
XX
DE Plasmid pBRV2 insert encoding granulocyte CSF.
XX
KW G-CSF; granulocyte colony stimulating factor; antimicrobial.
XX

Synthetic.

XX Key Location/Qualifiers
FH CDS 31..645
FT mat_peptide /tag= a
FT 121 /tag= b
XX
PN WO8604506-A.
PN WO8604605-A.
XX
PD 14-AUG-1986.
XX
XX 07-FEB-1986; 86WO-JP000053.
XX
XX 08-FEB-1985; 85JP-0023777.
PR 03-DEC-1985; 85JP-0270839.
XX
XX (CHUS) CHUGAI SEIYAKU KK.
PA (ONOM/) ONO M.
XX
XX Ono M, Nomura H, Tamura M, Matsumoto M;
XX
XX WPI; 1986-225384/34.
DR P-PSDB; AAP61341.
XX

XX Infection-protection inducing agent - contains human granulocyte
PT colony stimulating factor obtd. by cell culture or recombinant
PT technology.
XX

XX Disclosure; Fig 4; 139pp; Japanese.
XX

XX The plasmid encodes a novel polypeptide having human granulocyte
CC colony stimulating factor activity. The CSF may be produced either
CC from a cell-line secreting the factor into its medium, or from a
CC transformant microorganism. The product induces high immunity to

CC infection by anaerobic and aerobic bacteria and fungi.
XX
SQ Sequence 1521 BP; 304 A; 488 C; 402 G; 327 T; 0 other;

Query Match 93.3%; Score 486.8; DB 7; Length 1521;
Best Local Similarity 98.6%; Pred. No. 2.7e-101;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 25 CTACACAGAGCTTCCTTTTAAAAAGCTTAGAGCAAGTGAAGAAATCCAGGGCGATGCG 84
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
145 ctgccccagagcttctctcaagtcttagcagaagtgaagaaatccagggcgatggc 204
Qy 85 GCAGCGCTCCAGGAGAAGCTGTGCCACCTACAAAGCTGTGCCACCGAGAGAGCTGTG 144
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
205 gcagcgctcccaggagaagctgtgcccacctacaagctgtgccaccgcgaggagctggtg 264
Qy 145 CTGCTCGGACACTCTCTCGGGCATCCCTGGGCTCCCTTGAGCAGCTGCCAGCCAGGCC 204
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
265 ctgctcggacactctctgggcatccctgggctccctcctgagcagctgccccagcagcc 324
Qy 205 CTGCAGCTGGCAGGCTGCTTGAAGCAACTCCATAGCGGCTTTCCTTACACAGGGGCTC 264
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
325 ctgcagctggcagcgtgcttgagccaaactccatagcgcctttctctaccaggggttc 384
Qy 265 CTGCAGGCGCTGGAGAGGATCTCCCGAGTTGGGTCCACACTTGGACACACTGCAGCTG 324
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
385 ctgcagcgcctggaaggatctcccccgagttgggtccaccttggacacactgcagctg 444
Qy 325 GACGTGCGCGACTTTGGCCACCATCTGGCAGCAGATGGAGAACTGGGAATGCCCCCT 384
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
445 gacgtcgccgacttggccaccacatctggcagcagatgggaactgggaatggccct 504
Qy 385 GCCCTGACGACCCACCGAGGTGCCATGGCGCTTGCCTTCTGCTTCCAGCGCGGGCA 444
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
505 gccctgcagccccacaggtggcctgcccctcgcctctgcttccagcgccggca 564
Qy 445 GGAGGGTCTAGTTGCTCCACTCTGCAGAGCTTCCCTGGAGGTCTGTACCGGTTCTTA 504
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
565 ggaggggtcctagttgctctccatctgcagagcttctctggaggtgtcgtaccggttcta 624
Qy 505 CGCCACCTTGGCCAGGCC 522
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
625 cgcacacttggccagccc 642

RESULT 13

AAAN81478
ID AAAN81478 standard; cDNA; 1525 BP.
XX
AC AAAN81478;
XX
DT 14-DEC-1990 (first entry)
XX
DE Sequence encoding human granulocyte colony-stimulating factor (G-CSF)
DE in clone pP12 of MIA PaCa CDNA library.
XX

KW Leukaemia therapy; anaemia treatment; leukocyte; lymphokine;
KW interleukin-3; ss.
XX

OS Homo sapiens.
XX

FH Key Location/Qualifiers
FH CDS 43..132

FT mat_peptide /tag= a

FT 133..657 /tag= b

FT conflict /tag= c

FT /note="CHU-2 cDNA clones of Nagata et al. contains a
FT 9 base pair insertion (GTGATGGAG)"

FT conflict 588

FT /tag= d

```
FT conflict /note="G in the CHU-2 clone"
FT 1237 /*tag= e
FT /*tag="C in the CHU-2 clone"
PN EP256843-A.
XX 24-FEB-1988.
XX 11-AUG-1987; 87EP-0307114.
XX 18-NOV-1986; 86US-0932037.
PR 01-AUG-1986; 86US-0895194.
XX (CETU ) CETUS CORP.
XX Devlin JJ, Devlin PE, Kawasaki ES, Warren MK;
XX WPI; 1988-051585/08.
DR P-PSDB; AAP81162.
XX Human recombinant granulocyte colony-stimulating factor -
PT causes differentiation of granulocytes and monocytes from bone
PT marrow progenitor cells
XX Example; Fig 4; 61pp; English.
XX The major difference between this clone and the CHU-2 cDNA clones of
CC Nagata et al. (EMBO Journal 5:3,575-581(1986)) are given in FT of
CC AAN81478. The G-CSF was detected in purified mRNA fraction obtd. by
CC sucrose density gradient centrifugation of induced MIA PaCa-2 cells as
CC described in W08804607.
XX Sequence 1525 BP; 303 A; 491 C; 402 G; 329 T; 0 other;

Query Match 93.3%; Score 486.8; DB 9; Length 1525;
Best Local Similarity 98.6%; Pred. No. 2.7e-101;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 25 CTACACAGAGCTTCCCTTTTAAAGCTTAGAGCAAGTGAGGAGAGATCCAGGCGGATGGC 84
DB 157 ctgccacagagcttccctgctcaagtgtctagagcaagtggagagatccaggcgatggc 216
QY 85 GCAGCGCTCCAGGAGAGCTGTGTCACCTACAAGCTGTGCCACCCCGAGGAGCTGGTG 144
DB 217 gcagcgctccaggagaagctgtgcccacctaagaactgtgccaccccgaggagctgggtg 276
QY 145 CTGCTCGGACACTCTCTGGGATCCCTCTGGGCTCCCTTGAGCAGTGTGCCACCCAGGCC 204
DB 277 ctgctcggaactctctgggcatccctggctccctgagcagctgccccagccagggcc 336
QY 205 CTGCAGCTGGCAGGCTCTTGAGCACTCCATAGCGGCTTTTCTCTACAGGGGCTC 264
DB 337 ctgcagctggcagggctgttgagccaaactccatagggccttctctaccaggggctc 396
QY 265 CTGCAGGCTCTGGAAGGAGTCTCCCGAGTGTGGGTCCTCCACCTTGGACACACTGACGCTG 324
DB 397 ctgcagggcttggaaggaatctccccagctgggtccccacttggaacactgcaactg 456
QY 325 GACGTGCGGACTTTGGCAGACATCTGCGCAGCAGATGGAAGAACTGGGAATGGCCCT 384
DB 457 gacgtgcggaactttgcacacacatctggcagcagatggaagaactgggaatggccct 516
QY 385 GCCTTGAGCCACCCAGGAGTGCATGCCGCTTGCCTCTCTCTTCCAGCGCGGCA 444
DB 517 gccctgagccacacaggggtgcaatgcccgtcttgcctctgcttccagcgccgggca 576
QY 445 GGAGGGTCTCTAGTTCCTCCCATCTCTCAGAGCTTCTCTGAGAGTGTCTGACCGCTCTTA 504
DB 577 ggaaggggtcctagtgtcctcccatctgcagagcttctcctgaggtgctgaccgctcta 636
QY 505 CGCCACCTTGCCCGAGCCC 522
```

```
DB 637 cgccaccttgccagccc 654

RESULT 14
AAN91086
ID AAN91086 standard; DNA; 1525 BP.
XX AAN91086;
XX AC AAN91086;
XX 04-JUL-1990 (first entry)
XX Plasmid pP12 contg. colony stimulating factor-1 gene.
XX Colony stimulating factor-1; pP12; ss.
XX Homo sapiens.
XX OS
XX FH Key Location/Qualifiers
XX FT sig_peptide 43..132
XX FT /*tag= a
XX FT mat_peptide 133..654
XX FT /*tag= b
XX PN W08901038-A.
XX PD 09-FEB-1989.
XX 20-JUL-1988; 88WO-US02445.
XX 24-JUL-1987; 87US-0077188.
XX (CETU ) CETUS CORP.
XX PI Kawasaki ES, Devlin JJ, Martin G, O'Rourke E, Clark R;
XX WPI; 1989-061174/08.
DR P-PSDB; AAP95033.
XX Recombinant Bacovirus transfer vectors - used for prodn. of colony
PT stimulating factor in Baculovirus insect cell expression system.
XX Disclosure; Fig 3; 68pp; English.
XX The cDNA insert of pP12 contains 11 more bases than the CHU-2 G-CSF
CC clone. The major difference between this clone, derived from MIA
CC PaCa-2, and CHU-2 clone is a 9 bp insertion in the CHU-2 clone between
CC bases 237 and 238. There are two other differences: an A at position 588
CC (G in the CHU-2 clone) is a silent third base change, and a T at
CC position 1237(C in the CHU-2 clone) in the 3' untranslated region.
CC See also AAN91085.
XX Sequence 1525 BP; 303 A; 490 C; 403 G; 328 T; 1 other;

Query Match 93.3%; Score 486.8; DB 10; Length 1525;
Best Local Similarity 98.6%; Pred. No. 2.7e-101;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 25 CTACACACAGCTTCCCTTTTAAAGCTTAGAGCAAGTTAGAGGAGATCCAGGCGGATGGC 84
DB 157 ctgccacagagcttccctgctcaagtgtctagagcaagtggagagatccaggcgatggc 216
QY 85 GCAGGCTCTCCAGGAGAGCTGTGTCACCTACAAGCTGTGCCACCCCGAGGAGCTGGTG 144
DB 217 gcagcgctccaggagaagctgtgcccacctaagaactgtgccaccccgaggagctgggtg 276
QY 145 CTGCTCGGACACTCTCTGGGATCCCTCTGGGCTCCCTTGAGCAGTGTGCCACCCAGGCC 204
DB 277 ctgctcggaactctctgggcatccctggctccctgagcagctgccccagccagggcc 336
QY 205 CTGCAGCTGGCAGGCTCTTGAGCACTCCATAGCGGCTTTTCTCTACAGGGGCTC 264
```

Db 337 ctgcagctggcaggctgttgagccaaactccatagcggcctttctcttaccaggggtc 396
Qy 265 CTGAGGCCCTTGAAGGATCTCCCGAGTTGGGTCCACCTTGGACACACTGCGAGCTG 324
Db 397 ctgcagcccttgaaggatctcccagagttggtccacacttggacacactgcagctg 456
Qy 325 GACGTGCGCGACTTTGGCCACCATCTGCGCAGCAGATGGAAGAACTGGGAATGCCCT 384
Db 457 gacgtgcgcgactttgcccaccacctctggcagcagatggagaactgggaatggccct 516
Qy 385 GCCCTGAGCCACCCAGGNGCCATGCCGCTTCCGCTTCTGCTTTCAGCGCCGGSCA 444
Db 517 gccctgcagccaccacccagggctccatgcggccttcgctctgtttccagcgccggca 576
Qy 445 GGAGGGTCTTAGTTGGCTCCCATCTGCAGAGCTTCTGAGGTTCTGATCCGCGCTTCTA 504
Db 577 ggaaggtctcagttgctcccatctcagagcttctcagagttctcagaggtgtctacccggtcta 636
Qy 505 GCCCACCTTGGCCAGCCC 522
Db 637 cgcacacttggccagccc 654

RESULT 15

AAN71089
ID AAN71089 standard; DNA; 1415 BP.

XX AC AAN71089;

DT 26-APR-1991 (first entry)

XX DE Sequence of human granulocyte colony stimulating factor (hpG-CSF)
DE cDNA clone Ppo2.

XX KW Haematopoietic disorders; therapy; aplastic anaemia;
KW bone marrow transplant; burn wounds; leukaemia; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
FT CDS 1..36

FT /*tag= a

FT mat_peptide 37..561

FT /*tag= b

FT polyA_signal 1797..1802

FT /*tag= c

XX PN WO8701132-A.

XX PD 26-FEB-1987.

XX PF 22-AUG-1986; 86WO-U001708.

XX PR 03-MAR-1986; 86US-0835548.

XX PR 23-AUG-1985; 85US-0768959.

XX PR 23-AUG-1985; 85US-0768954.

XX PA (KIRI-) KIRIN-AMGEN INC.

XX PA (KIRI) KIRIN-AMGEN INC.

XX PI Souza LM;

XX DR WPI; 1987-064855/09.

XX DR P-PSDB; AAP70730.

XX PT Poly:peptide with granulocyte colony stimulating factor activity
PT - obtd. by recombinant DNA procedures for treating haematopoietic disorders

XX PS Disclosure; Page 22-24; 79pp; English.

XX CC The examples describe procedures for the designing of probes for
CC hpG-CSF cDNA and genomic clones, both of which are claimed.

CC Specifically claimed are DNA sequences encoding for (Ala 1)hpG-CSF;
CC (Ser 36, 42, 64 and 74)hpG-CSF and the corresponding Met-1 cpds.
CC The novelty is that hpG-CSF is the prod. of procarvotic or
CC eucaryotic expression of an exogenous DNA sequence. The construction
CC of hpG-CSF expression vectors is also described in the examples.

XX SQ Sequence 1415 BP; 265 A; 451 C; 380 G; 319 T; 0 other;

Query Match 93.0%; Score 485.2; DB 8; Length 1415;

Best Local Similarity 98.4%; Pred. No. 6.2e-101;

Matches 490; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 25 CTACCACAGAGCTTCTCTTTTAAAGCTTAGAGCAAGTGAGGAAGATCCAGGGCGATGCC 84
Db 61 ctgccccagagcttctctcaagtgttagagcaagtggagaagatccaggcgcatggc 120
Qy 85 GCAGGGCTCCAGGAGAAGCTGTGTGCCACCTACAGCTGTGCCACCCCGAGAGCTGGTG 144
Db 121 gcagcgctccaggagaagctgtgtgccacctacaagctgtgccaccccgaggagtggtg 180
Qy 145 CTGCTCGGACACTCTCTTGGCATCCCTGGGCTCCCTGAGCAGCTGCCACAGCCAGGCC 204
Db 181 ctgctcggaacactctctgggcatcccttggctcccttgagcagctgccccagcaggccc 240
Qy 205 CTGACAGCTGGCAGGCTGCTTGAGCCAACTCCATAGCGGCTTTTCTCTTACACAGGGGCTC 264
Db 241 ctgcagctggcagctgcttgagccaaactccatagcggcctttctctaccaggggtc 300
Qy 265 CTGACAGGCCCTGGAGAGGATCTCCCGAGTTGGTGGTCCACCTTGGACACACTGAGCTG 324
Db 301 ctgcaggccctggaaggatctccccagttgggtccacacctggacacactgcagctg 360
Qy 325 GACGTGCGCGACTTTTGGCCACCACTCTGGCAGCAGATGGGAAGAACTGGGAATGCCCT 384
Db 361 gacgtcgcgacttctgcccaccacctctgagcagcagatggaagaactgggaatggccct 420
Qy 385 GCCCTGACGCCACCCAGGGTGCCATGCCGCTTCGCTCTGTCTTTCAGCGCCGGSCA 444
Db 421 gccctgcagccaccacccagggtgcatgccccttcgctctgcttccagcgccggca 480
Qy 445 GGAGGGTCTTAGTTGGCTCCCATCTGCGAGCTTTCCTGAGGCTGCTGACCGCGTTCTA 504
Db 481 ggaagggctcctggttgcctcccatctgcagagcttctcctggaggtgctgtacgcggtcta 540
Qy 505 CGCCACCTTGGCCAGCCC 522
Db 541 cgcacacttggccagccc 558

Search completed: May 8, 2002, 12:28:35

Job time: 12566 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2002, 07:52:27 ; Search time 1794.34 Seconds
(without alignments)
3926.464 Million cell updates/sec

Title: US-09-680-514-4_COPY_526_1047
Perfect score: 522
Sequence: 1 GGCCCAACATCGCGCTC.....TACGCCACCTTGCCCGCC 522

Scoring table: IDENTITY NUC
Gapox 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: em_estba:**
- 2: em_esthum:**
- 3: em_estin:**
- 4: em_estmu:**
- 5: em_estov:**
- 6: em_estpl:**
- 7: em_estro:**
- 8: em_htc:**
- 9: gb_est1:**
- 10: gb_est2:**
- 11: gb_htc:**
- 12: gb_gss:**
- 13: em_gss_hum:**
- 14: em_gss_inv:**
- 15: em_gss_pln:**
- 16: em_gss_vrt:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	454.2	87.0	990	10	BM423896	BM423896 AGENCOURT
2	387.8	74.3	817	10	BM009358	BM009358 603629823
3	343.6	65.8	570	10	BI961242	BI961242 MONOL_7-G
4	323.6	62.0	548	10	BI960812	BI960812 MONOL_1-B
5	320.4	61.4	973	10	BI411128	BI411128 602962472
6	312.8	59.9	598	10	BG548320	BG548320 602575289
7	300.6	57.6	507	10	BE485194	BE485194 172130 BA
8	281.2	53.9	829	10	BM009247	BM009247 603629663
9	210.6	40.3	948	10	BI822673	BI822673 603036006
10	181	34.7	368	10	BI961002	BI961002 MONOL_3-H
11	150	28.7	509	10	BM256261	BM256261 518592 MA
12	135.8	26.0	327	10	BF848766	BF848766 QVO-EN010
13	67.6	13.0	673	10	BI961882	BI961882 MONOL_7-G
14	54.8	10.5	925	12	CNS0091P	AL053013 Drosophil
15	50.2	9.6	513	10	BE480590	BE480590 165716 BA
16	49	9.4	925	12	CNS0091P	AL053013 Drosophil
17	48.8	9.3	617	9	AI981598	AI981598 pat.pk006

c 18	43.2	8.3	461	10	BE598945	BE598945 P11_84_E1
c 19	43.2	8.3	559	10	BG556176	BG556176 EMI_68_F0
c 20	42.6	8.2	452	9	AA155632	AA155632 zo70d02_s
c 21	42.6	8.2	899	12	CNS01VYT	AL169742 Tetraodon
c 22	42.4	8.1	1203	12	CNS015Y4	AL106054 Drosophil
c 23	42.4	8.1	598	9	AA115932	AA115932 z101a10_s
c 24	41.4	7.9	421	9	AA058743	AA058743 zk70f07_s
c 25	41.4	7.9	478	10	BI345642	BI345642 374523 MA
c 26	41.2	7.9	676	9	BB647049	BB647049 BB647049
c 27	40.8	7.8	545	10	BF868325	BF868325 292084 BA
c 28	40.8	7.8	893	10	BM016196	BM016196 603642649
c 29	40.8	7.8	1011	10	BM474138	BM474138 AGENCOURT
c 30	40.6	7.8	447	10	BG604563	BG604563 WHE0947_C
c 31	40.6	7.8	845	9	AL572931	AL572931 AL572931
c 32	40.6	7.8	932	12	CNS0072Q	AL066742 Drosophil
c 33	40.6	7.8	1003	10	BI410408	BI410408 602964928
c 34	40.4	7.7	458	9	AW273202	AW273202 xr35c05_x
c 35	40.4	7.7	485	10	BE463718	BE463718 hx79f09_x
c 36	40.2	7.7	864	12	CNS022LE	AL178331 Tetraodon
c 37	40	7.7	368	10	BI961002	BI961002 MONOL_3-H
c 38	40	7.7	548	10	BI960812	BI960812 MONOL_1-B
c 39	40	7.7	570	10	BI961242	BI961242 MONOL_7-G
c 40	39.8	7.6	387	9	AW193322	AW193322 xl73e08_x
c 41	39.8	7.6	427	9	BE048584	BE048584 hr50h10_x
c 42	39.8	7.6	432	9	AW175624	AW175624 RC3-BT004
c 43	39.8	7.6	441	9	AI818468	AI818468 wk60c03_x
c 44	39.8	7.6	444	9	AW072844	AW072844 xa42q06_x
c 45	39.8	7.6	450	9	AW273147	AW273147 xr34c05_x

ALIGNMENTS

RESULT 1

BM423896

LOCUS AGENCOURT_6399123 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5517129

DEFINITION 5', mRNA sequence.

ACCESSION BM423896

VERSION BM423896.1 GI:18392108

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTF

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: L10CM2018 row: n column: 10

High quality sequence stop: 542.

Location/Qualifiers

1. 990

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5517129"

/clone_lib="NIH_MGC_41"

/tissue_type="amelanotic melanoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: skin; Vector: pOTB7; Site:1: XhoI; Site:2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California,

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

REFERENCE 1 (bases 1 to 570)
AUTHORS Vandenplas, M.L., Cordonnier-Pratt, M.-M., Sudman, M.L., Wentzel, V.E.,
Gingle, A.R., Pratt, L.H. and Moore, J.N.
TITLE An EST database from equine (Equus caballus) monocytes
JOURNAL Unpublished (2001)
COMMENT Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805

Email: mmpratt@uga.edu
Sequences have been trimmed to exclude polyA, vector and regions
below Phred quality 16. The threshold for high quality sequence is
20. Three-prime sequences, which are obtained with PolyMix or T7
sequencing primer, are presented as the reverse complement.

Seq primer: JEN REV

High quality sequence stop: 541

POLYA-No.

FEATURES
Source

Location/Qualifiers
1..570

/organism="Equus caballus"

/db_xref="taxon:9796"

/clone_lib="Monocytes (MONOL)"

/cell_type="isolated peripheral blood monocytes stimulated
with E. coli lipopolysaccharide"

/note="Vector: pBluescript SK(-) from Lambda ZapII;

Site_1: XhoI; Site_2: EcoRI; The library was made from
poly-A RNA in the cloning vector lambda ZAPII. Clones to
be sequenced were prepared by mass excision."

100 a 202 c 167 g 101 t

BASE COUNT
ORIGIN

Query Match 65.8%; Score 343.6; DB 10; Length 570;

Best Local Similarity 87.4%; Pred. No. 8.5e-68;

Matches 376; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Oy 25 CTACCACAGAGCTTCCTTTAAAGAGCTTAGAGCAAGTGAGGAAGATCCAGGCCGATGGC 84

Db 141 CTGCCCCAGAGCTTCCTCTCAAGTGTCTAGAGCAAGTGAGGAAGATCCAGGCCGATGGT 200

Oy 85 GCAGCGCTCCAGGAGAGCTGTGCCACCTACAGCTGTGCCACCCGAGGAGCTGGTG 144

Db 201 GCAGCGCTGCAGGACAGCTGTGTGCCACCCACAGCTGTGCCACCCCTCAGGAGCTCATG 260

Oy 145 CTGCTCGACACTCTCTCTGGGCATCCCTCGGGCTCCCTCTGAGCAGCTGCCCCAGCCAGGCC 204

Db 261 CTGCTGGGACACTCTCTGGGCATCCCTCGAGCTCTCTCTGAGCAGCTGTCTCCAGCCAGGCC 320

Oy 205 CTGAGCTGGCAGGCTGTGTAGCCCACTTATAGCGGCTTTTCTCTACCAAGGGGCTC 264

Db 321 CTGAGCTGGAGGGGTGCTGTGAGCCAACTCCACAGTGGCCCTCTCTCTACCAAGGGTCTC 380

Oy 265 CTGAGGCGCTTGGAGGAGTCTCCCGAGTTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 324

Db 381 CTGAGGCGCTTGGAGGAGTCTCCCGAGTTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 440

Oy 325 GACGTGCGCGACTTTGGCCACCACCATCTGGCAGCAGATGGGAAGACTGGGAATGGCCCT 384

Db 441 GACGTGCGCGACTTTGGCCACCACCATCTGGCAGCAGATGGGAAGACTGGGAATGGCCCT 500

Oy 385 CCCCTGAGAGCCACCGAGGTGCATGCGGGCTTTCGCTCTGCTTTCAGGCGGGGCA 444

Db 501 GTGTGACAGCCACCATGGGCCATGACCGACCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 560

Oy 445 GGAGGGTCC 454

Db 561 GGAGGGTGC 570.

RESULT 4
BI960812

LOCUS MONOL1_B08.bl_A005 Monocytes (MONOL) Equus caballus CDNA, mRNA

DEFINITION sequence.

ACCESSION BI960812

VERSION BI960812.1

KEYWORDS GI:16319015

SOURCE horse.

ORGANISM Equus caballus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

REFERENCE 1 (bases 1 to 548)

AUTHORS Vandenplas, M.L., Cordonnier-Pratt, M.-M., Sudman, M.L., Wentzel, V.E.,
Gingle, A.R., Pratt, L.H. and Moore, J.N.

TITLE An EST database from equine (Equus caballus) monocytes

JOURNAL Unpublished (2001)

COMMENT Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805

Email: mmpratt@uga.edu

Sequences have been trimmed to exclude polyA, vector and regions
below Phred quality 16. The threshold for high quality sequence is
20. Three-prime sequences, which are obtained with PolyMix or T7
sequencing primer, are presented as the reverse complement.

Seq primer: JEN REV

High quality sequence stop: 481

POLYA-No.

FEATURES
Source

Location/Qualifiers
1..548

/organism="Equus caballus"

/db_xref="taxon:9796"

/clone_lib="Monocytes (MONOL)"

/cell_type="isolated peripheral blood monocytes stimulated
with E. coli lipopolysaccharide"

/note="Vector: pBluescript SK(-) from Lambda ZapII;

Site_1: XhoI; Site_2: EcoRI; The library was made from
poly-A RNA in the cloning vector lambda ZAPII. Clones to
be sequenced were prepared by mass excision."

98 a 203 c 147 g 98 t

BASE COUNT
ORIGIN

Query Match 62.0%; Score 323.6; DB 10; Length 548;

Best Local Similarity 86.4%; Pred. No. 2.8e-63;

Matches 356; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Oy 25 CTACCACAGAGCTTCCTTTTAAAGAGCTTAGAGCAAGTGAGGAAGATCCAGGCCGATGGC 84

Db 137 CTGCCCCAGAGCTTCCTCTCAAGTGTCTAGAGCAAGTGAGGAAGATCCAGGCCGATGGT 196

Oy 85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAAGCTGTGCCACCCGAGGAGCTGGTG 144

Db 197 GCAGCGCTGCAGGACAGCTGTGTGCCACCCCAAGCTGTGCCACCCCTCAGGAGCTCATG 256

Oy 145 CTGCTCGACACTCTCTGGGCATCCCTCGGGCTCCCTTGAGCAGCTGCCCCAGCCAGGCC 204

Db 257 CTGCTGGGACACTCTCTGGGCATCCCTCGAGCTCTCTCTGAGCAGCTGTCCAGCCAGGCC 316

Oy 205 CTGAGCTGGGAGGCTGTGTAGCCCACTTCCATACGCGCTTTTCTCTTACCAAGGGCTC 264

Db 317 CTGAGCTGGAGGGGTGCTGTAGCCCACTTCCACAGTGGCTCTCTCTTACCAAGGGTCTC 376

Oy 265 CTGAGGCGCTTGGAGGAGTCTCCCGAGTTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 324

Db 377 CTGAGGCGCTTGGAGGAGTCTCCCGAGTTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 436

Oy 325 GACGTGCGCGACTTTGGCCACCACCATCTGGCAGCAGATGGGAAGACTGGGAATGGCCCT 384

Db 437 GACGTGCGCGACTTTGGCCACCACCATCTGGCAGCAGATGGGAAGACTGGGAATGGCCCT 496

```
QY 385 GCCCTGACGCCACCCAGGGTGCATGCCGCCCTTCCCTCTGCTTTCCAGC 436
Db 497 GTGGTGCACCCACCCATGGCCCATCGGACCTTCGCTCGGCTTCCANC 548

RESULT 5
BI411128 973 bp mRNA linear EST 14-AUG-2001
LOCUS 602962472P1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5118185 5',
DEFINITION mRNA sequence.
ACCESSION BI411128
VERSION BI411128.1 GI:15172051
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 973)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11287 row: o column: 18
High quality sequence start: 32
High quality sequence stop: 912.
FEATURES
source Location/Qualifiers
1..973
/organism="Mus musculus"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:5118185"
/clone_lib="NCI_CGAP_Lu33"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site.1: NotI; Site.2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5'
TGTACCAATCTGAAGTGGGAGCGCGCTCTGTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 192 a 295 c 254 g 232 t
ORIGIN
Query Match 61.4%; Score 320.4; DB 10; Length 973;
Best Local Similarity 77.1%; Pred. No. 1.7e-62;
Matches 390; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
QY 17 CCTCAGTCTACACAGAGCTTCCTTTTAAAGCTTAGAGCAAGTGAAGATCCAGG 76
Db 91 CCTCGCTCTCCCGGAGCTTCTGCTTAAGTCCCTGGAGCAAGTGAAGATCCAGG 150
QY 77 GCGATGGCGGAGCGCTCCAGAGAGTGTGTGCCACTACAGCTGTGCCACCCGAGG 136
Db 151 CCAGCGGCTGGTGTCTGTGAGCAGTGTGTGCCACTACAGCTGTGTCAACCCGAGG 210
QY 137 AGCTGGTGTCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTCAGCAGCTGCCCA 196
Db 211 AGCTGGTGTCTCGGCACTCTCTGGGGATCCCGAAGGCTTCCCTGAGTGTCTCTA 270
```

```
Query Match          59.9%; Score 312.8; DB 10; Length 598;
Best Local Similarity 99.4%; Pred. No. 7.9e-61;
Matches 314; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 207 GCAGCTGCAGGCTGCTTTGAGCAACTCATAGCGGCTTTCCCTCTACCAAGGGCTCTCT 266
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 185 GAAGCTGGCAGGCTGCTTTGAGCAACTCATAGCGGCTTTCCCTCTACCAAGGGCTCTCT 244
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 267 GCAGGCTTGAAGGATCTCCCGAGTGGGTCACCTTGGACACTGGACACTGCAGCTGA 326
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 245 GCAGGCTTGAAGGATCTCCCGAGTGGGTCACCTTGGACACTGGACACTGCAGCTGA 304
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 327 CGTCCCGACTTTGCCACCAACCATCTGCAGCAGATGGAAGAACTGGGAATGGCCCTTGC 386
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 305 CGTCCCGACTTTGCCACCAACCATCTGCAGCAGATGGAAGAACTGGGAATGGCCCTTGC 364
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 387 CCTGAGCCCAACCCAGGTCATGCCGCGCTTCCCTCTCTGCTTTCCAGCGCCGGCAGG 446
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 365 CCTGAGCCCAACCCAGGTCATGCCGCGCTTCCCTCTCTGCTTTCCAGCGCCGGCAGG 424
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 447 AGGGTCTAGTTCCTCCCATCTGCAGAGCTTCTTGAGGTGCTGATCCGGTCTTACG 506
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 425 AGGGTCTAGTTCCTCCCATCTGCAGAGCTTCTTGAGGTGCTGATCCGGTCTTACG 484
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 507 CCACCTTCCCGAGCCC 522
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 485 CCACCTTCCCGAGCCC 500
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
LOCUS BE485194 507 bp mRNA linear EST 28-AUG-2000
DEFINITION 172130 BARC 5BOV Bos taurus cdna 5', mRNA sequence.
ACCESSION BE485194
VERSION BE485194.1 GI:9604727
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
          Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 507)
AUTHORS Sonstegard,T.S., Capuco,A.V., Van Tassel,C.P., Ashwell,M.S. and
          Wells,K.D.
TITLE Mapping of Expressed Sequence Tags from a normalized bovine mammary
          gland cDNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Sonstegard TS
          USDA, ARS, Beltsville Agricultural Research Center
          Bldg. 200 Rm 2A, Beltsville, MD 20705, USA
          Tel: 301 504 8416
          Fax: 301 504 8414
          Email: tadselpsi@ars.usda.gov
          Single pass sequencing. Bases called and alt_trimmed with phred
          v0.980904.e. Vector identified by cross_match with the "minscore 18
          and -minmatch 12 options.
          PCR Primers
          FORWARD: AGGAACAGCTATGACCAT
          BACKWARD: GTTTCCTAGTCAGACG
          Plate: 134 row: J column: 17
          Seq primer: ATTTAGGTACACTATAG.
          Location/Qualifiers
            1..507
              /organism="Bos taurus"
              /db_xref="taxon:9913"
              /clone_lib="BARC 5BOV"
              /tissue_type="pooled"
              /lab_host="DH10B"
              /note="vector: pCMV SPORT6; Site.1: XbaI; Site.2: XhoI;
              Library made from pooled mRNA isolated from mammary
              tissues at eight physiological, developmental, and disease
              states."

Query Match          57.6%; Score 300.6; DB 10; Length 507;
Best Local Similarity 84.1%; Pred. No. 4.3e-58;
Matches 339; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Qy 105 GGTGCGCACTTACAGCTGTGCCACCCGAGGAGCTGGTGTCTCGGACACTCTCTGGG 164
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 104 GGTGCGCGCCCAACAAGCTGTGCCACCCGAGGAGCTGATGCTGCTCAGCACTCTCTGGG 163
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 165 CATCCCGCTGGGCTCCCTCGAGCAGCTGCCACGAGGCGCTTCAGCTGGCAGGCTGCTT 224
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 164 CATCCCGCAGGCTCCCTTAAGCAGCTGTCTCCAGCCAGTCCCTGCAGCTGACGAGCTGCCT 223
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 225 GAGCCAACTCCATAGCGGCTTTTCTCTACAGGGGCTCTTCGAGGGCTCGAAGGAT 284
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 224 GAACCAACTACAGCGCGGCTCTTTCTCTACAGGGCTCTTCGAGGGCTCGCGGGCAT 283
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 285 CTCGCCCGAGTTGGCTCCACCTTGGACACACTGCAGCTGGAGCTCGCGGACTTTGCCAC 344
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 284 CTCGCCAGAGCTGGCCCGCCACCTTGGACACACTGCAGCTGGAGCTGACACTGACTTTGCCAC 343
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 345 CACCATCTGGCAGCAGATGGAAACTGGGAATGGCCCTGCGCTTCGAGCCACCCAGG 404
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 344 GAACATCTGGCTGCAGATGGAGGACCTGGGGCGGCCCGCTGTGCAGCCACCCAGGG 403
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 405 TGCCATGGCGGCTTCGCTCTGCTTTCAGCGCGGGGAGGAGGGTCTTCTAGTTGCCTC 464
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 404 CGCCATGGCGGCTTCGCTCTGCTTTCAGCGCGGGGAGGAGGGTCTTCTAGTTGCCTC 463
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 465 CCATCTGCAGAGCTTCCTGGAGGTGCTGCTACCGCTTCTACGC 507
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 464 CCAGCTGCATGCTTTCTTCTGGAGCTGGCATACCGTGGCCTGCSC 506
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
LOCUS BM009247 829 bp mRNA linear EST 30-OCT-2001
DEFINITION 603629663F1 NIH_MGC_41 Homo sapiens cdna clone IMAGE:5434958 5',
          mRNA sequence.
ACCESSION BM009247
VERSION BM009247.1 GI:16523601
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 829)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: DCTD/BTP
          cDNA Library Preparation: Ling Hong/Rubin Laboratory
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLCM1912 row: n column: 15
          High quality sequence start: 7
          High quality sequence stop: 590.
          Location/Qualifiers
            1..829
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="IMAGE:5434958"
              /clone_lib="NIH_MGC_41"
              /tissue_type="amelanotic melanoma, cell line"
              /lab_host="DH10B (phage-resistant)"

FEATURES
          source
```


[illegible]

```

DEFINITION OVO-EN0102-081100-458-g06 EN0102 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF848766
VERSION BF848766.1 GI:12235903
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 327)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV0&t2=QV0-EN0102-
081100-458-g06&t3=2000-11-08&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 327.
FEATURES
source .
1..327
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0102"
/dev_stage="Adult"
/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 61 a 87 c 103 g 76 t
ORIGIN
Query Match 26.0%; Score 135.8; DB 10; Length 327;
Best Local Similarity 89.6%; Pred. No. 7e-21;
Matches 146; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 201 GGCCCTGCAGCTGCAGGCTGTTGAGCCAACTCCATAGCGGCTTTCTCTACACAGG 260
Db 139 GGCCCATACACAGGCGCGCTGTTTGAGCAACTCCATAGCGGCTTTCTCTACACAGG 198
QY 261 GCTCTGCGACCCCTGGAGGATCTCCCGGAGTTGGTCCACCTTGGACACTGCA 320
Db 199 GCTCTGCGACCCCTGGAGGATCTCCCGGAGTTGGTCCACCTTGGACACTGCA 258
QY 321 GCTGACGTCGCCGACTTTGGCCACCATCTGGCAGCAGATG 363
Db 259 GCTGACGTCGCCGACTTTGGCCACCATCTGGCAGCAGTG 301

RESULT 13
LOCUS BI961882
DEFINITION MON01_7_g12_g1_A005 Monocytes (MON01) Equus caballus cDNA, mRNA
sequence.
ACCESSION BI961882

```

```

VERSION BI961882.1 GI:16320085
KEYWORDS EST.
SOURCE horse.
ORGANISM Equus caballus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Perissodactyla; Equidae; Equus.
1 (bases 1 to 673)
Vandenplas, M.L., Cordonnier-Pratt, M.-M., Sudman, M.L., Wentzel, V.E.,
Gingle, A.R., Pratt, L.H. and Moore, J.N.
TITLE An EST database from equine (Equus caballus) monocytes
JOURNAL Unpublished (2001)
COMMENT Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below phred quality 16. The threshold for high quality sequence is
20. Three-prime sequences, which are obtained with Polymix or T7
sequencing primer, are presented as the reverse complement.
Seq primer: Polymix
High quality sequence start: 39
High quality sequence stop: 673
POLYA=No.
FEATURES
source .
1..673
Location/Qualifiers
/organism="Equus caballus"
/db_xref="taxon:9796"
/clone_lib="Monocytes (MON01)"
/cell_type="Isolated peripheral
with E. coli lipopolysaccharide"
/note="Vector: pBluescript SK(-) from Lambda ZapII;
Site_1: XhoI; Site_2: EcoRI; The library was made from
poly-A RNA in the cloning vector lambda ZapII. Clones to
be sequenced were prepared by mass excision."
BASE COUNT 147 a 178 c 166 t
ORIGIN
Query Match 13.0%; Score 67.6; DB 10; Length 673;
Best Local Similarity 83.5%; Pred. No. 2.2e-05;
Matches 76; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 432 CCAGCGCCGGGAGGAGGGTCTAGTTCCTCCACTGCGAGCTTCTCTGGAGGTCTC 491
Db 1 CCAGCGCCGGGAGGAGGGTCTAGTTCCTCCACTGCGAGCTTCTCTGGAGGTCTC 60

QY 492 GTACCGCGTTTCTACGCCACCTTGCCAGCCC 522
Db 61 ATACCGTGGTCTGCGCTACCTGCGGAGCCC 91

RESULT 14
LOCUS CNS0091P
DEFINITION Drosophila melanogaster genome survey sequence TPT3 end of BAC #
BACR19D16 of RPI-98 library from Drosophila melanogaster (fruit
fly); genomic survey sequence.
ACCESSION AL053013
VERSION AL053013.1 GI:4934461
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
REFERENCE Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
AUTHORS Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 925)
Genoscope.
Direct Submission
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

```

us-09-680-514-4_copy_526_1047.ige

Wed May 8 14:05:50 2002

GenCore version 4.5

Copyright (c) 1993 - 2000 CompuGen Ltd.

Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
 May 8, 2002, 07:53:17 ; Search time 1915.42 seconds
 (without alignments)
 5703.006 Million cell updates/sec

Run on:

US-09-680-514-4_COPY_526_1047

Title:

Perfect score:

Sequence:

Scoring table:

Searched:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.pat.*

6: gb.ph.*

7: gb.pl.*

8: gb.pr.*

9: gb.ro.*

10: gb.sts.*

11: gb.sy.*

12: gb.un.*

13: gb.vi.*

14: em.ba.*

15: em.fun.*

16: em.in.*

17: em.in.*

18: em.mu.*

19: em.om.*

20: em.or.*

21: em.ov.*

22: em.pat.*

23: em.ph.*

24: em.pl.*

25: em.ro.*

26: em.sts.*

27: em.un.*

28: em.vi.*

29: em.htg.hum.*

30: em.htg.inv.*

31: em.htg.other.*

32: em.htgo.inv.*

33: em.htgo.inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Score Match Length DB ID Description

Result Score Match Length DB ID

Result Score Match Length DB ID

Result Score Match Length DB ID

Result Score Match Length DB ID

Result Score Match Length DB ID

Result Score Match Length DB ID

Result Score Match Length DB ID

Result Score Match Length DB ID

1	520.4	99.7	6	E02574
2	514.4	98.5	6	E02575
3	512.4	98.2	6	AR024358
4	486.8	93.3	6	AR091731
5	486.8	93.3	6	E01731
6	486.8	93.3	6	E02573
7	486.8	93.3	6	E01716
8	486.8	93.3	6	E15131
9	486.8	93.3	6	E15131
10	486.8	93.3	6	E15131
11	486.8	93.3	6	E15131
12	486.8	93.3	6	E15131
13	486.8	93.3	6	E15131
14	486.8	93.3	6	E15131
15	486.8	93.3	6	E15131
16	486.8	93.3	6	E15131
17	486.8	93.3	6	E15131
18	486.8	93.3	6	E15131
19	486.8	93.3	6	E15131
20	486.8	93.3	6	E15131
21	486.8	93.3	6	E15131
22	486.8	93.3	6	E15131
23	486.8	93.3	6	E15131
24	486.8	93.3	6	E15131
25	486.8	93.3	6	E15131
26	486.8	93.3	6	E15131
27	486.8	93.3	6	E15131
28	486.8	93.3	6	E15131
29	486.8	93.3	6	E15131
30	486.8	93.3	6	E15131
31	486.8	93.3	6	E15131
32	486.8	93.3	6	E15131
33	486.8	93.3	6	E15131
34	486.8	93.3	6	E15131
35	486.8	93.3	6	E15131
36	486.8	93.3	6	E15131
37	486.8	93.3	6	E15131
38	486.8	93.3	6	E15131
39	486.8	93.3	6	E15131
40	486.8	93.3	6	E15131
41	486.8	93.3	6	E15131
42	486.8	93.3	6	E15131
43	486.8	93.3	6	E15131
44	486.8	93.3	6	E15131
45	486.8	93.3	6	E15131

ALIGNMENTS

525 bp DNA linear PAT 29-SEP-1997
 525 bp DNA linear PAT 29-SEP-1997
 525 bp DNA linear PAT 29-SEP-1997

RESULT 1

E02574

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

E02574

DNA encoding human colony-stimulating factor derivative.

E02574

E02574.1

GI:2170804

JP 199027075-A/2

synthetic construct

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

E02574

DNA encoding human colony-stimulating factor derivative.

E02574

E02574.1

GI:2170804

JP 199027075-A/2

synthetic construct

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

E02574

DNA encoding human colony-stimulating factor derivative.

E02574

E02574.1

GI:2170804

JP 199027075-A/2

synthetic construct

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

E02574

DNA encoding human colony-stimulating factor derivative.

E02574

E02574.1

GI:2170804

JP 199027075-A/2

synthetic construct

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

E02574

DNA encoding human colony-stimulating factor derivative.

E02574

E02574.1

GI:2170804

JP 199027075-A/2

synthetic construct

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

artificial sequence

E02574

DNA encoding human colony-stimulating factor derivative.

E02574

E02574.1

GI:2170804

JP 199027075-A/2

synthetic construct

artificial sequence

artificial sequence

artificial sequence

artificial sequence

PI ITOU SEIGA
PC C12N9/72,C07K13/00,C07K15/14,C12N1/21,C12N5/10,C12N9/64, PC
C12N15/27,
PC C12N15/58,C12N15/70,C12N15/85,C12P21/02,C12P21/02,(C12N1/21,
C12R1:19),
PC (C12P21/02,C12R1:19),(C12P21/02,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH key
FT mat_peptide 1..522
 /product='human colony-stimulating factor FT
 derivative which
 is named 'hg-CSF[ND28]';
FT Location/Qualifiers
FEATURES
source
1..525
/organism='synthetic construct'
/db_xref='taxon:32630'
BASE COUNT 94 a 184 c 148 g 99 t
ORIGIN
Query Match 99.7%; Score 520.4; DB 6; Length 525;
Best Local Similarity 99.8%; Pred. No. 8.8e-87;
Matches 521; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCGCAACATATCGCGCTCGAGTCTACACAGAGCTTCTTTTAAAGAGCTTAGAGCAA 60
DB 1 GCACCAACATATCGCGCTCGAGTCTACACAGAGCTTCTTTTAAAGAGCTTAGAGCAA 60
QY 61 GTGAGGAAGATCCAGGCGATGCGCAGCGCTCCAGGAGAAGCTGTGCCACCTACAAG 120
DB 61 GTGAGGAAGATCCAGGCGATGCGCAGCGCTCCAGGAGAAGCTGTGCCACCTACAAG 120
QY 121 CTGTGCCACCCGAGGAGCTGTGTCTCGGACACTCTTGGGCATCCCTGGGCTCCC 180
DB 121 CTGTGCCACCCGAGGAGCTGTGTCTCGGACACTCTTGGGCATCCCTGGGCTCCC 180
QY 181 CTGAGCAGCTGCCCGCAGCGCTTGCAGCTGGCAGCTGTGAGCCAACTCCATAGC 240
DB 181 CTGAGCAGCTGCCCGCAGCGCTTGCAGCTGGCAGCTGTGAGCCAACTCCATAGC 240
QY 241 GGCCTTTTCTTCTACAGGGGCTCTGAGGGCTTGAAGGATCTCCCGAGTTGGGT 300
DB 241 GGCCTTTTCTTCTACAGGGGCTCTGAGGGCTTGAAGGATCTCCCGAGTTGGGT 300
QY 301 CCCACCTTGGACACTGACAGCTGGAGCTGCGGACTTTGCCACCACTCTTGGCAGCAG 360
DB 301 CCCACCTTGGACACTGACAGCTGGAGCTGCGGACTTTGCCACCACTCTTGGCAGCAG 360
QY 361 ATGGAAGAACTGGGAATGGCCCTGCTGCAGCCACCGAGGCTGCCATGCCGCTTC 420
DB 361 ATGGAAGAACTGGGAATGGCCCTGCTGCAGCCACCGAGGCTGCCATGCCGCTTC 420
QY 421 GCCTCTGCTTCCAGCGCGGAGAGGGTCTTCTAGTCCCTCCATCTGCGAGCTTC 480
DB 421 GCCTCTGCTTCCAGCGCGGAGAGGGTCTTCTAGTCCCTCCATCTGCGAGCTTC 480
QY 481 CTGAGGCTGCTACCGGCTTCTACGCCACTTGGCCAGGCC 522
DB 481 CTGAGGCTGCTACCGGCTTCTACGCCACTTGGCCAGGCC 522
RESULT 2
LOCUS E02575 525 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding human colony-stimulating factor derivative.
ACCESSION E02575
VERSION E02575.1 GI:2170805
KEYWORDS JP 1990227075-A/3.
SOURCE synthetic construct.

ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 525)
AUTHORS Sasaki,K., Nishi,T., Yasumura,S., Sato,M. and Ito,S..
TITLE NEW POLYPEPTIDE
JOURNAL Patent: JP 1990227075-A 3 10-SEP-1990;
KYOWA HAKKO KOGYO CO LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1990227075-A/3
PD 10-SEP-1990
PF 28-SEP-1989 JP 1989253097
PR 29-SEP-1988 JP 88P 245705
PI SASAKI KATSUTOSHI, NISHI TATSUYA, YASUMURA SHIGEYOSHI, PI
SAITO MORIYUKI,
PI ITOU SEIGA
PC C12N9/72,C07K13/00,C07K15/14,C12N1/21,C12N5/10,C12N9/64, PC
C12N15/27,
PC C12N15/58,C12N15/70,C12N15/85,C12P21/02,C12P21/02,(C12N1/21,
C12R1:19),
PC (C12P21/02,C12R1:19),(C12P21/02,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH key
FT mat_peptide 1..522
 /product='human colony-stimulating factor FT
 derivative(hg-CSF[ND28N6]) which is resistant FT
 to proteases'.
FT Location/Qualifiers
FEATURES
source
1..525
/organism='synthetic construct'
/db_xref='taxon:32630'
BASE COUNT 94 a 183 c 148 g 100 t
ORIGIN
Query Match 98.5%; Score 514; DB 6; Length 525;
Best Local Similarity 99.0%; Pred. No. 1.3e-85;
Matches 517; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 GCGCAACATATCGCGCTCGAGTCTTACCACAGAGCTTCTTTTAAAGAGCTTAGAGCAA 60
DB 1 GCACCTACATATCGCAACTCGAGTCTTACCACAGAGCTTCTTTTAAAGAGCTTAGAGCAA 60
QY 61 GTGAGGAAGATCCAGGCGATGCGCAGCGCTCCAGGAGAAGCTGTGCCACCTACAAG 120
DB 61 GTGAGGAAGATCCAGGCGATGCGCAGCGCTCCAGGAGAAGCTGTGCCACCTACAAG 120
QY 121 CTGTGCCACCCGAGGAGCTGTGTCTCGGACACTCTTGGGCATCCCTGGGCTCCC 180
DB 121 CTGTGCCACCCGAGGAGCTGTGTCTCGGACACTCTTGGGCATCCCTGGGCTCCC 180
QY 181 CTGAGCAGCTGCCCGCAGCGCTTGCAGCTGGCAGCTGTGAGCCAACTCCATAGC 240
DB 181 CTGAGCAGCTGCCCGCAGCGCTTGCAGCTGGCAGCTGTGAGCCAACTCCATAGC 240
QY 241 GGCCTTTTCTTCTACAGGGGCTCTGAGGGCTTGAAGGATCTCCCGAGTTGGGT 300
DB 241 GGCCTTTTCTTCTACAGGGGCTCTGAGGGCTTGAAGGATCTCCCGAGTTGGGT 300
QY 301 CCCACCTTGGACACTGACAGCTGGAGCTGCGGACTTTGCCACCACTCTTGGCAGCAG 360
DB 301 CCCACCTTGGACACTGACAGCTGGAGCTGCGGACTTTGCCACCACTCTTGGCAGCAG 360
QY 361 ATGGAAGAACTGGGAATGGCCCTGCTGCAGCCACCGAGGCTGCCATGCCGCTTC 420
DB 361 ATGGAAGAACTGGGAATGGCCCTGCTGCAGCCACCGAGGCTGCCATGCCGCTTC 420
QY 421 GCCTCTGCTTCCAGCGCGGAGAGGGTCTTCTAGTCCCTCCATCTGCGAGCTTC 480
DB 421 GCCTCTGCTTCCAGCGCGGAGAGGGTCTTCTAGTCCCTCCATCTGCGAGCTTC 480

Db 421 GCCTCTGCTTCCAGCGCGGACAGAGGGTCTAGTTGCTCCACTCGAGAGCTTC 480

Qy 481 CTGAGGCTGCTACCGGTTCTAGCGCCACCTTGCACAGCC 522

Db 481 CTGAGGCTGCTACCGGTTCTAGCGCCACCTTGCACAGCC 522

RESULT 3

E02576 525 bp DNA linear PAT 29-SEP-1997

LOCUS E02576

DEFINITION DNA encoding human colony-stimulating factor derivative.

ACCESSION E02576

VERSION E02576.1 GI:2170806

KEYWORDS JP 1990227075-A/4.

SOURCE synthetic construct.

ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 525)

AUTHORS Sasaki,K., Nishi,T., Yasumura,S., Sato,M. and Ito,S.

TITLE NEW POLYPEPTIDE

JOURNAL Patent: JP 1990227075-A 4 10-SEP-1990;

COMMENT KYOWA HAKKO KOGYO CO LTD

OS Artificial gene

OC Artificial sequence; Genes.

PN JP 1990227075-A/4

PD 10-SEP-1990

PF 28-SEP-1989 JP 1989253097

PI 29-SEP-1988 JP 88P 245705

PI SASAKI KATSUTOSHI, NISHI TATSUYA, YASUMURA SHIGEYOSHI, PI

SATO MORIYUKI,

PI ITOU SEIGA

PC C12N9/72,C07K13/00,C07K15/14,C12N1/21,C12N5/10,C12N9/64, PC

C12N15/27,

PC C12N15/58,C12N15/70,C12N15/85,C12P21/02,C12P21/02,(C12N1/21,

C12P1/19)

PC (C12P21/02,C12P1/19),(C12P21/02,C12P1/91);

CC strandedness: Double;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

Key Location/Qualifiers

FT mat_peptide 1..522

FT /product="human colony-stimulating factor FT

FT derivative

FT (hg-CSF[ND28N145]) which is resistant to FT

FT proteases;

FEATURES

source Location/Qualifiers

1..525

/organism="synthetic construct"

/db_xref="taxon:32630"

BASE COUNT 96 a 182 c 146 g 101 t

ORIGIN

Query Match 98.2%; Score 512.4; DB 6; Length 525;

Best Local Similarity 98.9%; Pred. No. 2.6e-85;

Matches 516; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GCGCAACATATCGCGCTCGAGTCTACACAGAGCTTCCTTTAAAGCTTAGAGCAA 60

Db 1 GCACCAACATATCGCGCTCGAGTCTACACAGAGCTTCCTTTAAAGCTTAGAGCAA 60

Qy 61 GTGAGGAAGATCCAGGCGATGGCGAGCGCTCCAGAGAGCTGTGTGCCACCTACAAG 120

Db 61 GTGAGGAAGATCCAGGCGATGGCGAGCGCTCCAGAGAGCTGTGTGCCACCTACAAG 120

Qy 121 CTGTGCCACCCGAGGAGCTGGTCTCTCGGACACACTCTCTGGGCATCCCTGGGCTCCC 180

Db 121 CTGTGCCACCCGAGGAGCTGGTCTCTCGGACACACTCTCTGGGCATCCCTGGGCTCCC 180

Qy 181 CTGAGGAGCTGCCCGACCGAGCGCTCTGACGCTGTGTGAGCCAACTGCATPAGC 240

Db 181 CTGAGGAGCTGCCCGACGAGCGCTCGAGCTGGCAGGCTGCTTGCAGCAACTCCATAGC 240

Qy 241 GGCCTTTTCTCTTACCAGGGGCTCTGAGGCGCTTGAAGGAGTCTCCCCGAGTTGGGT 300

Db 241 GGCCTTTTCTCTTACCAGGGGCTCTGAGGCGCTTGAAGGAGTCTCCCCGAGTTGGGT 300

Qy 301 CCCACCTTGGACACACTGACGCTGACGCTGCCGACTTTTGCACACACCATCTGCAGCAG 360

Db 301 CCCACCTTGGACACACTGACGCTGACGCTGCCGACTTTTGCACACACCATCTGCAGCAG 360

Qy 361 ATGAAGAAGCTGGAATGGCCCTGCGCTGCGAGCCACCCAGGTTGCCATCTGCAGCAG 420

Db 361 ATGAAGAAGCTGGAATGGCCCTGCGCTGCGAGCCACCCAGGTTGCCATCTGCAGCAG 420

Qy 421 GCCTCTGCTTCCAGCGCGGAGGAGGGTCTAGTTGCCTCCCATCTGCAGAGCTTC 480

Db 421 GCCTCTGCTTCCAGCGCGGAGGAGGGTCTAGTTGCCTCCCATCTGCAGAGCTTC 480

Qy 481 CTGAGGCTGCTGCTACCGGCTTCTACGCGACCTTGCACAGCC 522

Db 481 CTGAGGCTGCTGCTACCGGCTTCTACGCGACCTTGCACAGCC 522

RESULT 4

AR024358

LOCUS AR024358

DEFINITION Sequence 1 from patent US 5795968.

ACCESSION AR024358

VERSION AR024358.1 GI:3977652

KEYWORDS Unknown.

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 525)

AUTHORS Kuga,T., Miyaj,H., Sato,M., Okabe,M., Morimoto,M., Itoh,S.,

Yamasaki,M., Yokoo,Y., Yamaguchi,K., Yoshida,H. and Komatsu,Y.

TITLE Polypeptide derivatives of human granulocyte colony stimulating factor

JOURNAL Patent: US 5795968-A 1 18-AUG-1998;

FEATURES

source Location/Qualifiers

1..525

/organism="unknown"

BASE COUNT 85 a 192 c 151 g 97 t

ORIGIN

Query Match 93.3%; Score 486.8; DB 6; Length 525;

Best Local Similarity 98.6%; Pred. No. 1.4e-80;

Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 25 CTACACAGAGCTTCCTTTTAAAGCTTAGAGCAAGCTTAGAGAGATCCAGGGCGATGGC 84

Db 25 CTGCCCCAGAGCTTCCTGCTCAAGTGTAGAGCAAGTGAAGAGATCCAGGGCGATGGC 84

Qy 85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTTACAAGCTGTGCCACCCCGAGGAGCTGTG 144

Db 85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTTACAAGCTGTGCCACCCCGAGGAGCTGTG 144

Qy 145 CTGCTCGGACACTCTCTGGGATCCCTCGGCTCCCTGAGCAGCTGCCCGAGCAGGCC 204

Db 145 CTGCTCGGACACTCTCTGGGATCCCTCGGCTCCCTGAGCAGCTGCCCGAGCAGGCC 204

Qy 205 CTGAGCTGCGAGGCTCTGTGAGCAACTCCATAGCGGCTTTTCTTACCGGGGCTC 264

Db 205 CTGAGCTGCGAGGCTCTGTGAGCAACTCCATAGCGGCTTTTCTTACCGGGGCTC 264

Qy 265 CTGAGCGCTTGAAGGAGTCTCCCGAGTTGGGTGCCACCTTGGACACACTGCAGCTG 324

Db 265 CTGAGCGCTTGAAGGAGTCTCCCGAGTTGGGTGCCACCTTGGACACACTGCAGCTG 324

Qy 325 GAGCTGCGGACCTTTGCCACCCACTCTGCGACAGATGGAAGAACTGGGAATGGCCCT 384

Db 325 GAGCTGCGGACCTTTGCCACCCACTCTGCGACAGATGGAAGAACTGGGAATGGCCCT 384

QY 385 GCCTGAGCCACCCAGGTCATCGCGGCTTGCCCTGCTCTTCCAGGCGCGGCA 444
|||||
Db 385 GCCTGAGCCACCCAGGTCATCGCGGCTTGCCCTGCTCTTCCAGGCGCGGCA 444
|||||
QY 445 GGAGGGTCTAGTTGGCTCCATCTGCAGAGCTTCTCGAGGTGTCGACCGGTTCTA 504
|||||
Db 445 GGAGGGTCTAGTTGGCTCCATCTGCAGAGCTTCTCGAGGTGTCGACCGGTTCTA 504
|||||
QY 505 CGCCACCTTGCCAGGCC 522
|||||
Db 505 CGCCACCTTGCCAGGCC 522
|||||
RESULT 5
LOCUS AR091731 525 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 1 from patent US 5994518.
ACCESSION AR091731
VERSION AR091731.1 GI:10018485
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 525)
AUTHORS Kuga,T., Miyaji,H., Sato,M., Okabe,M., Morimoto,M., Itoh,S.,
Yamasaki,M., Yokoo,Y., Yamaguchi,K., Yoshida,H. and Komatsu,Y.
TITLE Method of producing a polypeptide having human granulocyte colony
stimulating factor activity
JOURNAL Patent: US 5994518-A 1 30-NOV-1999;
FEATURES
source Location/Qualifiers
1..525
/organism="unknown"
BASE COUNT 85 a 192 c 151 g 97 t
ORIGIN

Query Match 93.3%; Score 486.8; DB 6; Length 525;
Best Local Similarity 98.6%; Pred. No. 1.4e-80;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 25 CTACCACAGAGCTCTTTTAAAGCTTAGAGCAAGTTCAGGAGATCCAGGCGGATGCC 84
|||||
Db 25 CTGCCCCAGAGCTCTCTGCTAAGTCTTAGAGCAAGTTCAGGAGATCCAGGCGGATGCC 84
|||||
QY 85 GCAGCGCTCCAGGAGAGCTGTGTGCACCTACAAGCTGTGCCACCCCGAGGAGCTGGTG 144
|||||
Db 85 GCAGCGCTCCAGGAGAGCTGTGTGCACCTACAAGCTGTGCCACCCCGAGGAGCTGGTG 144
|||||
QY 145 CTGCTGGGACACTCTCTGGGCAATCCCTTGGGCTTCCCTGAGCAGCTGCCCCAGCGGCC 204
|||||
Db 145 CTGCTGGGACACTCTCTGGGCAATCCCTTGGGCTTCCCTGAGCAGCTGCCCCAGCGGCC 204
|||||
QY 205 CTCAGCTGGCAGGCTCTTGGCCAACTCCATAGCGGCTTTCTCTACCCAGGGGCTC 264
|||||
Db 205 CTCAGCTGGCAGGCTCTTGGCCAACTCCATAGCGGCTTTCTCTACCCAGGGGCTC 264
|||||
QY 265 CTCAGCGCTCGGAAGGATCTCCCGAGTTGGGTGCCACCTTGGACACACTGCAGCTG 324
|||||
Db 265 CTCAGCGCTCGGAAGGATCTCCCGAGTTGGGTGCCACCTTGGACACACTGCAGCTG 324
|||||
QY 325 GAGCTGCGGACTTTGCCACCACTATCTGGCAGAGATGGAAGAACTGGGAATGCCCT 384
|||||
Db 325 GAGCTGCGGACTTTGCCACCACTATCTGGCAGAGATGGAAGAACTGGGAATGCCCT 384
|||||
QY 385 GCCTCGAGCCACCCAGGTCATCGCGGCTTGCCCTGCTCTGCTTTCCAGGCGCGGCA 444
|||||
Db 385 GCCTCGAGCCACCCAGGTCATCGCGGCTTGCCCTGCTCTGCTTTCCAGGCGCGGCA 444
|||||
QY 445 GGAGGGTCTAGTTGCCTCCATCTGCAGAGCTTCTCGAGGTGTCGACCGGTTCTA 504
|||||
Db 445 GGAGGGTCTAGTTGCCTCCATCTGCAGAGCTTCTCGAGGTGTCGACCGGTTCTA 504
|||||

QY 505 CGCCACCTTGCCAGGCC 522
|||||
Db 505 CGCCACCTTGCCAGGCC 522
|||||
RESULT 6
LOCUS E01731 525 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA encoding hg-CSF form human peripheral macrophage.
ACCESSION E01731
VERSION E01731.1 GI:2169984
KEYWORDS JP 1988267292-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 525)
AUTHORS Kuga,T., Komatsu,Y., Miyaji,H., Sato,M., Okabe,M., Morimoto,M.,
Itoh,S., Yamazaki,M., Yokoo,Y. and Yamaguchi,K.
TITLE NOVEL POLYPEPTIDE
JOURNAL Patent: JP 1988267292-A 1 04-NOV-1988;
COMMENT KYOWA HAKKO KOGYO CO LTD
OS Human
PN JP 1988267292-A/1
PD 04-NOV-1988
PF 23-DEC-1987 JP 1987326384
PR 23-DEC-1986 JP 86P 306799
PI KUGA TEISUO, KOMATSU YUKI, MIYAJI HIROMASA, SATO MORIYUKI, PI
OKABE MASAMI,
PI MORIMOTO MAKOTO, ITOU SEIGA, YAMAZAKI MOTOO, YOKOO YOSHIHARU,
PI YAMAGUCHI KAZUO
PC C12P21/02,C07K13/00,C12N1/20,C12N15/00//A61K37/02,A61K37/02,
PC A61K37/02,
PC (C12P21/02,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC *source: cell_line=Periferal macrophage;
FH Key Location/Qualifiers
FT mat_peptide 1..525
FT /product="mature peptide of hg-CSF".
FEATURES
source Location/Qualifiers
1..525
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 85 a 192 c 151 g 97 t
ORIGIN

Query Match 93.3%; Score 486.8; DB 6; Length 525;
Best Local Similarity 98.6%; Pred. No. 1.4e-80;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 25 CTACCACAGAGCTCTTTTAAAGCTTAGAGCAAGTTCAGGAGATCCAGGCGGATGCC 84
|||||
Db 25 CTGCCCCAGAGCTCTCTGCTAAGTCTTAGAGCAAGTTCAGGAGATCCAGGCGGATGCC 84
|||||
QY 85 GCAGCGCTCCAGGAGAGCTGTGTGCACCTACAAGCTGTGCCACCCCGAGGAGCTGGTG 144
|||||
Db 85 GCAGCGCTCCAGGAGAGCTGTGTGCACCTACAAGCTGTGCCACCCCGAGGAGCTGGTG 144
|||||
QY 145 CTGCTCGGACACTCTCTGGGCAATCCCTTGGGCTTCCCTGAGCAGCTGCCCCAGCGGCC 204
|||||
Db 145 CTGCTCGGACACTCTCTGGGCAATCCCTTGGGCTTCCCTGAGCAGCTGCCCCAGCGGCC 204
|||||
QY 205 CTGCACTGGCAGGCTGCTTAGGCCAACTCCATAGCGGCTTTCTCTACCCAGGGGCTC 264
|||||
Db 205 CTGCACTGGCAGGCTGCTTAGGCCAACTCCATAGCGGCTTTCTCTACCCAGGGGCTC 264
|||||
QY 265 CTGCAAGGCTCGGAAGGATCTCCCGAGTTGGGTGCCACCTTGGACACACTGCAGCTG 324
|||||
Db 265 CTGCAAGGCTCGGAAGGATCTCCCGAGTTGGGTGCCACCTTGGACACACTGCAGCTG 324
|||||


```

FT CDS 1..525
FT /product='human G-CSF'.
FEATURES
  source
    Location/Qualifiers
      1..525
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
BASE COUNT 85 a 192 c 151 g 97 t
ORIGIN

Query Match 93.3%; Score 486.8; DB 6; Length 525;
Best Local Similarity 98.6%; Pred. No. 1.4e-80;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 25 CTACACAGAGCTTCCTTTTAAAGAGCTTAGACAGAGTGGAGAGATCCAGGCGGATGC 84
  |||
DB 25 CTGCCCCAGAGCTTCCTCAAGTCTTAGACAGAGTGGAGAGATCCAGGCGGATGC 84
  |||
QY 85 GCAGCGCTCCAGGAGAGCTGTGTGCACCTACAAGCTGTGCCACCCAGGAGAGCTGGT 144
  |||
DB 85 GCAGCGCTCCAGGAGAGCTGTGTGCACCTACAAGCTGTGCCACCCAGGAGAGCTGGT 144
  |||
QY 145 CTGCTCGGACACTCTCTGGGATCCCTTGGCTCCCTGAGCAGCTGCCACCCAGGAGGCC 204
  |||
DB 145 CTGCTCGGACACTCTCTGGGATCCCTTGGCTCCCTGAGCAGCTGCCACCCAGGAGGCC 204
  |||
QY 205 CTGACAGTGGCAGAGCTGTGAGCCAACTCCATAGCGGCTTTTCTCTACACAGGCGTC 264
  |||
DB 205 CTGACAGTGGCAGAGCTGTGAGCCAACTCCATAGCGGCTTTTCTCTACACAGGCGTC 264
  |||
QY 265 CTGACAGTGGCAGAGCTGTGAGCCAACTCCATAGCGGCTTTTCTCTACACAGGCGTC 264
  |||
DB 265 CTGACAGTGGCAGAGCTGTGAGCCAACTCCATAGCGGCTTTTCTCTACACAGGCGTC 264
  |||
QY 325 GACGTGCGGAGCTTTGCCACACCATCTGGCAGCAGATGGAGAACTGGGAATGGCCCT 384
  |||
DB 325 GACGTGCGGAGCTTTGCCACACCATCTGGCAGCAGATGGAGAACTGGGAATGGCCCT 384
  |||
QY 385 GACGTGCGGAGCTTTGCCACACCATCTGGCAGCAGATGGAGAACTGGGAATGGCCCT 384
  |||
DB 385 GACGTGCGGAGCTTTGCCACACCATCTGGCAGCAGATGGAGAACTGGGAATGGCCCT 384
  |||
QY 445 GGAGGGGTCTAGTTCCTCCATCTGCAGAGCTTCCCTGGAGGTGTCGTACCGGCTCTA 504
  |||
DB 445 GGAGGGGTCTAGTTCCTCCATCTGCAGAGCTTCCCTGGAGGTGTCGTACCGGCTCTA 504
  |||
QY 505 CGCCACCTTGCCAGGCC 522
  |||
DB 505 CGCCACCTTGCCAGGCC 522
  |||

RESULT 10
LOCUS I71150
DEFINITION Sequence 1 from patent US 5681720.
ACCESSION I71150
VERSION I71150.1
KEYWORDS GI:3007285
SOURCE Unknown.
ORGANISM Unknown.

PI KUGA TETSUO, KOMATSU YUKI, MIYAJI HIROMASA, SATO MORIYUKI, PI
ORABE MASAMI.
PI MORIMOTO MAKOTO, ITOU SEIGA, YAMAZAKI MOTOO, YOKOO YOSHIIHARU,
PI YAMAGUCHI KAZUO
PC C12N15/09,A61K38/00,C07K14/535,C12N1/21,C12P21/02,(C12P21/02,
PC C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC key Location/Qualifiers
FH
FH
FT source 1..525
FT /organism="Homo sapiens"
FT /cell_type="peripheral blood macrophage" FT
mat_peptide 1..522
FT /product="G-CSF".
FEATURES
  source
    Location/Qualifiers
      1..525
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
BASE COUNT 85 a 192 c 151 g 97 t
ORIGIN

Query Match 93.3%; Score 486.8; DB 6; Length 525;
Best Local Similarity 98.6%; Pred. No. 1.4e-80;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 25 CTACACAGAGCTTCCTTTTAAAGAGCTTAGACAGAGTGGAGAGATCCAGGCGGATGC 84
  |||
DB 25 CTGCCCCAGAGCTTCCTCAAGTCTTAGACAGAGTGGAGAGATCCAGGCGGATGC 84
  |||
QY 85 GCAGCGCTCCAGGAGAGCTGTGTGCACCTACAAGCTGTGCCACCCAGGAGAGCTGGT 144
  |||
DB 85 GCAGCGCTCCAGGAGAGCTGTGTGCACCTACAAGCTGTGCCACCCAGGAGAGCTGGT 144
  |||
QY 145 CTGCTCGGACACTCTCTGGGATCCCTTGGCTCCCTGAGCAGCTGCCACCCAGGAGGCC 204
  |||
DB 145 CTGCTCGGACACTCTCTGGGATCCCTTGGCTCCCTGAGCAGCTGCCACCCAGGAGGCC 204
  |||
QY 205 CTGACAGTGGCAGAGCTGTGAGCCAACTCCATAGCGGCTTTTCTCTACACAGGCGTC 264
  |||
DB 205 CTGACAGTGGCAGAGCTGTGAGCCAACTCCATAGCGGCTTTTCTCTACACAGGCGTC 264
  |||
QY 265 CTGACAGTGGCAGAGCTGTGAGCCAACTCCATAGCGGCTTTTCTCTACACAGGCGTC 324
  |||
DB 265 CTGACAGTGGCAGAGCTGTGAGCCAACTCCATAGCGGCTTTTCTCTACACAGGCGTC 324
  |||
QY 325 GACGTGCGGAGCTTTGCCACACCATCTGGCAGCAGATGGAGAACTGGGAATGGCCCT 384
  |||
DB 325 GACGTGCGGAGCTTTGCCACACCATCTGGCAGCAGATGGAGAACTGGGAATGGCCCT 384
  |||
QY 385 GACGTGCGGAGCTTTGCCACACCATCTGGCAGCAGATGGAGAACTGGGAATGGCCCT 444
  |||
DB 385 GACGTGCGGAGCTTTGCCACACCATCTGGCAGCAGATGGAGAACTGGGAATGGCCCT 444
  |||
QY 445 GGAGGGGTCTAGTTCCTCCATCTGCAGAGCTTCCCTGGAGGTGTCGTACCGGCTCTA 504
  |||
DB 445 GGAGGGGTCTAGTTCCTCCATCTGCAGAGCTTCCCTGGAGGTGTCGTACCGGCTCTA 504
  |||
QY 505 CGCCACCTTGCCAGGCC 522
  |||
DB 505 CGCCACCTTGCCAGGCC 522
  |||

RESULT 9
LOCUS E15131
DEFINITION Human mRNA for granulocyte-colony stimulating factor.
ACCESSION E15131
VERSION E15131.1
KEYWORDS GI:5709814
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 525)
Kuga,T., Komatsu,Y., Miyaji,H., Sato,M., Okabe,M., Morimoto,M.,
Itou,S., Yamazaki,M., Yokoo,Y. and Yamaguchi,K.
NEW POLYPEPTIDE
Patent: JP 1998052281-A 1 24-FEB-1998;
KYOWA HAKKO KOGYO CO LTD
OS Homo sapiens (human)
PN JP 1998052281-A/1
PD 24-FEB-1998
PF 23-DEC-1987 JP 1997114630
PR 23-DEC-1986 JP 86P 306799
```



```
XX OS Homo sapiens (Human)
CC PN JP 1995149798-A/1
CC PD 13-JUN-1995
CC PF 08-AUG-1994 JP 1994185787
CC PR 23-DEC-1986 JP 86P 306799
CC PI KUGA TETSUO, KOMATSU YUKI, MIYAJI HIROMASA, SATO MORIYUKI,
CC PI OKABE MASAMI,
CC PI MORIMOTO MAKOTO, ITOU SEIGA, YAMAZAKI MOTOO, YOKOO YOSHIHARU,
CC PI YAMAGUCHI KAZUO
CC PC C07K14/535, C12N1/21, C12N15/09, C12P21/02, A61K38/00, C12N1/21,
CC PC C12R1/19,
CC PC C12P21/02, C12R1/19;
CC CC strandedness: Single;
CC CC topology: Linear;
CC FH Key Location/Qualifiers
CC FT source 1..525
CC FT /organism="Homo sapiens"
XX Key Location/Qualifiers
FH source 1..525
FH /db_xref="taxon:9606"
FT /organism="Homo sapiens"
FT
XX Sequence 525 BP; 85 A; 192 C; 151 G; 97 T; 0 other;

Query Match 93.3%; Score 486.8; DB 23; Length 525;
Best Local Similarity 98.6%; Pred. No. 1.4e-80;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 25 CTACCAGAGCTTCCTTTTAAAGCTTAGACCAAGTGGAGGAAGTCCAGGCGGATGCC 84
DB 25 CTGCCCCAGAGCTTCCTGCTCAAGTCTTAGACCAAGTGGAGGAAGTCCAGGCGGATGCC 84
QY 85 GCAGCGCTCCAGAGAGCTGTGCGCACCTACAAGCTGTGCCACCCCGAGGAGCTGGTG 144
DB 85 GCAGCGCTCCAGAGAGCTGTGCGCACCTACAAGCTGTGCCACCCCGAGGAGCTGGTG 144
QY 145 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTCAGCAGCTGCCCCAGCGAGGCC 204
DB 145 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTCAGCAGCTGCCCCAGCGAGGCC 204
QY 205 CTGAGCTGGCAGGCTGTGAGCCAACTCCATAGCGGCTTTTCTCTACACAGGGGCTC 264
DB 205 CTGAGCTGGCAGGCTGTGAGCCAACTCCATAGCGGCTTTTCTCTACACAGGGGCTC 264
QY 265 CTCAGGCCCTGGAAGGATCTCCCGGAGTTGGGTCCACCTTGGACACACTGCAGCTG 324
DB 265 CTCAGGCCCTGGAAGGATCTCCCGGAGTTGGGTCCACCTTGGACACACTGCAGCTG 324
QY 325 GAGCTCGCGACACTCTGCGCACCACTCTGCGAGAGTGGGAAGTGGGAATGCCCTT 384
DB 325 GAGCTCGCGACACTCTGCGCACCACTCTGCGAGAGTGGGAAGTGGGAATGCCCTT 384
QY 385 GCCCTGACGCCACCCAGGGTGCCATGCGGCTTTCGCTCTGCTTTCCAGCGCGGCA 444
DB 385 GCCCTGACGCCACCCAGGGTGCCATGCGGCTTTCGCTCTGCTTTCCAGCGCGGCA 444
QY 445 GGAGGGTCTTAGTTCCTCCATCTGCAGAGCTTCCTGGAGGTGCTGACGGCGTTCTA 504
DB 445 GGAGGGTCTTAGTTCCTCCATCTGCAGAGCTTCCTGGAGGTGCTGACGGCGTTCTA 504
QY 505 CGCCACTTGGCCAGGCC 522
DB 505 CGCCACTTGGCCAGGCC 522

RESULT 13
HUMGCSF HUMGCSF 660 bp mRNA linear PRI 27-APR-1993
LOCUS
```

DEFINITION

Human granulocyte colony stimulating factor mRNA, complete cds.

ACCESSION

M17706

VERSION

M17706.1 GI:183040

KEYWORDS

granulocyte colony stimulating factor.

SOURCE

Human MIA Paca-2 cell line, cDNA to mRNA, (library of Kawasaki et al.), clone pP12.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 660)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

Devlin, J.J., Devlin, P.E., Myambo, K., Lilly, M.B., Rado, T.A. and Warren, M.K.

TITLE

Expression of granulocyte colony-stimulating factor by human cell lines

JOURNAL

J. Leukoc. Biol. 41, 302-306 (1987)

MEDLINE

87196936

FEATURES

source

Location/Qualifiers

1..660

/organism="Homo sapiens"

/db_xref="taxon:9606"

43..132

/note="granulocyte colony stimulating factor signal peptide"

43..657

/note="granulocyte colony stimulating factor precursor"

/codon_start=1

/protein_id="AAA35882.1"

/db_xref="GI:183041"

/translation="MAGPATOSPMKLMALQLLWHSALTVOENTPLGPASSLPQSFL

LKCLEQVRKIQDGNALQKLCNTYKLCHPPELVLLGSLGIPWAPLSSCPQALQIA

GCLQSLGSLFLYGLLQALGELISPELGTLDLQDVADFPATTIQQWELGMAPAL

QPTQGMFAFAFQRRAGVFLVASHLQSFLEVSRYLRHLAQP"

133..654

/product="granulocyte colony-stimulating factor"

BASE COUNT 116 a 243 c 188 g 113 t

ORIGIN Unreported.

Query Match

93.3%; Score 486.8; DB 9; Length 660;

Best Local Similarity

98.6%; Pred. No. 1.4e-80;

Matches

491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 25 CTACCAGAGCTTCCTTTTAAAGCTTAGACCAAGTGGAGGAAGTCCAGGCGGATGCC 84

DB 157 CTGCCCCAGAGCTTCCTGCTCAAGTCTTAGACCAAGTGGAGGAAGTCCAGGCGGATGCC 216

QY 85 GCAGCGCTCCAGAGAGCTGTGCGCACCTACAAGCTGTGCCACCCCGAGGAGCTGGTG 144

DB 217 GCAGCGCTCCAGAGAGCTGTGCGCACCTACAAGCTGTGCCACCCCGAGGAGCTGGTG 276

QY 145 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTCAGCAGCTGCCCCAGCGAGGCC 204

DB 277 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTCAGCAGCTGCCCCAGCGAGGCC 336

QY 205 CTGAGCTGGCAGGCTGTGAGCCAACTCCATAGCGGCTTTTCTCTACACAGGGGCTC 264

DB 337 CTGAGCTGGCAGGCTGTGAGCCAACTCCATAGCGGCTTTTCTCTACACAGGGGCTC 396

QY 265 CTGAGGCCCTGGAAGGATCTCCCGGAGTTGGGTCCACCTTGGACACACTGCAGCTG 324

DB 397 CTGAGGCCCTGGAAGGATCTCCCGGAGTTGGGTCCACCTTGGACACACTGCAGCTG 456

QY 325 GAGCTCGCGACTTTGCCACCACTCTGGCAGAGTGGGAAGTGGGAATGGCCCTT 384

DB 457 GAGCTCGCGACTTTGCCACCACTCTGGCAGAGTGGGAAGTGGGAATGGCCCTT 516

QY 395 GCCCTGACGCCACCCAGGGTGCCATGCGGCTTTCGCTCTGCTTTCCAGCGCGGCA 444

DB 517 GCCCTGACGCCACCCAGGGTGCCATGCGGCTTTCGCTCTGCTTTCCAGCGCGGCA 576

QY 445 GGAGGGTCTTAGTTCCTCCATCTGCAGAGCTTCCTGGAGGTGCTGACGGCGTTCTA 504

DB 577 GGAGGGTCTTAGTTCCTCCATCTGCAGAGCTTCCTGGAGGTGCTGACGGCGTTCTA 636

QY 505 CGCCACCTTGCCAGGCC 522
|||||
Db 637 CGCCACCTTGCCAGGCC 654

RESULT 14
HSCGSR1
LOCUS Human mRNA for granulocyte colony-stimulating factor (G-CSF)
(PBRV-2).
ACCESSION X03655
VERSION X03655.1 GI:31693
KEYWORDS colony stimulating factor; signal peptide.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1498)
AUTHORS Nagata,S., Tsuchiya,M., Asano,S., Yamamoto,O., Hirata,Y.,
Kubota,N., Oheda,M., Nomura,H. and Yamazaki,T.
TITLE The chromosomal gene structure and two mRNAs for human granulocyte
colony-stimulating factor.
JOURNAL EMBO J. 5 (3), 575-581 (1986)
MEDLINE 86220137
COMMENT Data kindly reviewed (19-JUN-1986) by S. Nagata.
FEATURES
Location/Qualifiers
source
1..1498
/organism="Homo sapiens"
/db_xref="taxon:9606"
31..120
31..645
/codon_start=1
/product="G-CSF protein"
/protein_id="CAA27290.1"
/db_xref="GI:31693"
/translation="MAGPATQSPMKLMALQLLIHWSALMTVOEATPLGPASSLPQSFL
LKLEOVRIQGDGAALQKLCATYKLCHEPELVLLGSLGIPWPLSPSCPSQALOLA
GCLSLHSLFLYQGLQALEGTSPELGTLDLTLDVADFATTIQWQEEELGMAL
QPTQGAMPAPASAFQRRAGGLVLAHSLQFLEVSRYRLHQAQ"
121..642
mat_peptide
/product="mature G-CSF protein (aa 1-174)"
1481..1486
misc_feature
/note="polyadenylation signal"
1498
polyA_site
/note="polyadenylation site"
BASE COUNT 281 a 488 c 402 g 327 t
ORIGIN

Query Match 93.3%; Score 486.8; DB 9; Length 1498;
Best Local Similarity 98.6%; Pred. No. 1.2e-80;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 25 CTACACAGAGCTTCCTTTTAAAGCTTAGACAAAGTAGGAGATCCAGGCGATGGC 84
|||
Db 145 CTGCCCCAGAGCTTCTGCTCAAGTGCTTAGAGCAAGTAGGAGATCCAGGCGATGGC 204
|||
QY 85 GCAGCGCTCCAGGAGAGCTGTGCGCACCTACAGCTGTGCCACCCAGGAGCTGGT 144
|||
Db 205 GCAGCGCTCCAGGAGAGCTGTGCGCACCTACAGCTGTGCCACCCAGGAGCTGGT 264
|||
QY 145 CTGCTGGACACACTCTGGGCATCCCTGGCTCCCTGAGCAGCTGCCACCCAGGCGCC 204
|||
Db 265 CTGCTGGACACACTCTGGGCATCCCTGGCTCCCTGAGCAGCTGCCACCCAGGCGCC 324
|||
QY 205 CTGAGCTGGCAGGCTGCTTTAGCCCACTCATAGCGGCTTTTCTCTACCGGGGCTC 264
|||
Db 325 CTGAGCTGGCAGGCTGCTTTAGCCCACTCATAGCGGCTTTTCTCTACCGGGGCTC 384
|||
QY 265 CTGAGCGCTTGAGGAGTCTCCCGGAGTTGGGTGCCACCTTGGACACACTGACGCTG 324
|||

Db 385 CTGACGCGCTTGAAGGATCTCCCGAGTTGGTCCACCTTGGACACTGCAGCTG 444
|||||
QY 325 GACGTGCGCGACTTTTGCACACCATCTGGCAGCAGATGGAAGACTGGGAATGCCCT 384
|||||
Db 445 GACGTGCGCGACTTTTGCACACCATCTGGCAGCAGATGGAAGACTGGGAATGCCCT 504
|||||
QY 385 GCCCTGCAGGCCACCCAGGTTGCCATGCCGCGCTTCCCTCTGCTTCCAGCGCGGGCA 444
|||||
Db 505 GCCCTGCAGGCCACCCAGGTTGCCATGCCGCGCTTCCCTCTGCTTCCAGCGCGGGCA 564
|||||
QY 445 GGAGGGTCTAGTTGCTCCATCTGCAGAGCTTCTCGAGGTGTCGTACCGGTTCTA 504
|||||
Db 565 GGAGGGTCTAGTTGCTCCATCTGCAGAGCTTCTCGAGGTGTCGTACCGGTTCTA 624
|||||
QY 505 CGCCACCTTGCCAGGCC 522
|||||
Db 625 CGCCACCTTGCCAGGCC 642

RESULT 15
E01219
LOCUS cDNA encoding human G-CSF.
DEFINITION E01219
ACCESSION E01219
VERSION E01219.1 GI:2169478
KEYWORDS JP 1987129298-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1521)
AUTHORS Yamazaki,T., Yamamoto,O., Hirata,Y., Sekimori,Y. and Osada,J.
TITLE NOVEL POLYPEPTIDE
JOURNAL Patent: JP 1987129298-A 1 11-JUN-1987;
CHUGAI PHARMACEUT CO LTD
COMMENT
OS Human
PN JP 1987129298-A/1
PD 11-JUN-1987
PF 02-DEC-1985 JP 1985269455
PI YAMAZAKI TATSUMI, YAMAMOTO OSAMI, HIRATA YUICHI, PI SEKIMORI
YASUO,
FI OSADA JUICHI
PC C07K13/00,C07H21/04,C12N15/00,C12P21/02;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: cell_type=oral cavity tumor cells;
CC *source: cell_line=CHU-2;
FH Key Location/Qualifiers
FT CDS 31..645
FT /product="human G-CSF"
FT sig_peptide 31..120
FT /note="human G-CSF prewso"
FT mat_peptide 121..642
FT /note="human G-CSF"
FT 5'UTR 1..30
FT 3'UTR 645..1499
FT polyA_site 1500..1522
FT mRNA 1..1522.
FEATURES
Location/Qualifiers
source
1..1521
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 304 a 489 c 401 g 327 t
ORIGIN

Query Match 93.3%; Score 486.8; DB 6; Length 1521;
Best Local Similarity 98.6%; Pred. No. 1.2e-80;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy 25 CTACCAGAGCTTCCTTTTAAAGCTTAGACCAAGTGAGGAAGATCCAGGGCGATGGC 84
Db 145 CTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAAGTGAGGAAGATCCAGGGCGATGGC 204
Qy 85 GCAGCGCTCCAGGAGAGAGCTGTGTGCCACCTACAAGCTGTGCCACCCGAGAGAGCTGGTG 144
Db 205 GCAGGGCTCCAGGAGAGAGCTGTGTGCCACCTACAAGCTGTGCCACCCGAGAGAGCTGGTG 264
Qy 145 CTGCTCGGACACTCTCTGGGCATGCCCTGGGCTCCCTTGAGCAGCTGCCCCAGCCAGGCC 204
Db 265 CTGCTCGGACACTCTCTGGGCATGCCCTGGGCTCCCTTGAGCAGCTGCCCCAGCCAGGCC 324
Qy 205 CTGCAGCTGGCAGGCTGCTTGAGCCAACTCCATAGGGGCCCTTTTCTCTACCAAGGGCTC 264
Db 325 CTGCAGCTGGCAGGCTGCTTGAGCCAACTCCATAGGGGCCCTTTTCTCTACCAAGGGCTC 384
Qy 265 CTGCAGGGCCCTGGAGGGATCTCCCCGAGTTGGGTCCACACTTGGACACACTGCAGCTG 324
Db 385 CTGCAGGGCCCTGGAGGGATCTCCCCGAGTTGGGTCCACACTTGGACACACTGCAGCTG 444
Qy 325 GACGTCGCCGACTTGGCCACCACTATGCGACAGATGGAAGAACTGGGAATGGCCCT 384
Db 445 GACGTCGCCGACTTGGCCACCACTATGCGACAGATGGAAGAACTGGGAATGGCCCT 504
Qy 385 GCCCTGCAGCCCAACCCAGGGTGCCATGCCGGCTTCGCCCTCTGCTTCCAGCCCGGGCA 444
Db 505 GCCCTGCAGCCCAACCCAGGGTGCCATGCCGGCTTCGCCCTCTGCTTCCAGCCCGGGCA 564
Qy 445 GGAGGGTCTCTAGTTGGCTCCCATCTGCAGAGCTTCTGAGAGTGTCGTACCGCTTCTA 504
Db 565 GGAGGGTCTCTAGTTGGCTCCCATCTGCAGAGCTTCTGAGAGTGTCGTACCGCTTCTA 624
Qy 505 CGCCACCTTGCCCGAGCCC 522
Db 625 CGCCACCTTGCCCGAGCCC 642
```

Search completed: May 8, 2002, 10:02:06
Job time: 7729 sec